Constructivist Instruction

Success or Failure?

Edited by

Sigmund Tobias

Institute for Urban and Minority Education, Teachers College, Columbia University

Thomas M. Duffy

School of Education, Indiana University

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2009
by Routledge
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Typeset in Minion by Wearset Ltd, Boldon, Tyne and Wear
Printed and bound in the United States of America on acid-free paper by
Edwards Brothers, Inc.

Library of Congress Cataloging-in-Publication Data
Constructivist instruction : success or failure? / [edited by]
Sigmund Tobias, Thomas M. Duffy.
p. cm.
1. Constructivism (Education) 2. Instructional systems. I. Tobias,
Sigmund. II. Duffy, Thomas M.
LB1590.3.C6785 2009
370.15'2–dc22 2009005021

ISBN10: 0-415-99423-3 (hbk)
ISBN10: 0-415-99424-1 (pbk)
ISBN10: 0-203-87884-1 (ebk)

ISBN13: 978-0-415-99423-1 (hbk)
ISBN13: 978-0-415-99424-8 (pbk)
ISBN13: 978-0-203-87884-2 (ebk)

For Lora and Cindy

For Sam & Dan
From one [illegible] to
another with best
wishes

Sig

5 Beyond More Versus Less

A Reframing of the Debate on Instructional Guidance

Alyssa Friend Wise and Kevin O'Neill

Simon Fraser University

In this chapter we attempt to reframe the debate surrounding instructional guidance in a way that may be more productive than the one pursued recently. Up till now, the conversation between constructivists and instructionists has largely centered on the adversarial question of whether or not constructivist instructional approaches provide enough guidance to be effective. However, we argue that experimental "high versus low guidance" studies cannot provide a valid basis for making inferences about the fundamental merits of constructivist teaching. Reviewing some of the literature cited in the recent debate (with a particular focus on worked-example studies), we will argue that for constructivists and instructionists alike, the *quantity* of guidance is just one dimension along which guidance can be usefully characterized. We introduce the *context* in which guidance is delivered and the *timing* with which guidance is delivered as two more important concerns. We then make a case for a research agenda that we believe may bring constructivists and instructionists together in exploring questions about the optimal quantity, context, and timing of guidance in ill-defined problem domains.

On one side of the recent debate is a perspective that sees instructional approaches such as problem-based learning (PBL) and guided inquiry as providing too little guidance to support learning effectively (e.g., Kirschner, Sweller, & Clark, 2006). Proponents of this view often cite lab-based findings such as the worked-example effect (e.g., Sweller & Cooper, 1985) as support for the position that "more" guidance, generally given at the outset of instruction, is almost always best for learners. Because of the primacy this perspective places on the delivery of up-front explicit instruction, we refer to it as "instructionism."

On the other side of the debate, there is a perspective that claims that inquiry and PBL approaches do provide a great deal of guidance, and may in fact produce learning outcomes that are superior to up-front explanation (e.g., Hmelo-Silver, Duncan, & Chinn, 2007). Proponents of this view often cite classroom-based studies showing that students in inquiry-based classrooms achieve greater depth of understanding than students in traditional ones (e.g., Hickey, Kindfield, Horwitz, & Christie, 1999) as support for this position, generally referred to as "constructivism."

It appears to us that the two sides are talking past each other. We suspect that this may be due to different ideas about the purpose of guidance, different

ambitions with regard to the transfer of learning, and different views about the nature of the evidence needed to justify claims about the merits of *any* instructional approach. In the following sections, we try to unpack these differences and show how these incommensurate views lead us to a methodological catch-22 that, at present, makes the debate irresolvable.

The reader may wonder, if we cannot answer the question of which approach is better, is there still a way to move forward productively? We present one possible path that uses the amount, context, and timing of guidance as a toolkit for thinking about the design of instruction. We explore evidence from both the constructivist and instructionist perspectives regarding the nature of effective instructional guidance, and attempt to develop principles that both instructionists and constructivists may be able to agree to. By using a common language, we hope to engender the possibility of research agendas of interest to both instructionists and constructivists, and to which each group can contribute.

Transfer of Learning and the Purpose of Guidance: Contrasting Instructionism and Constructivism

A logical place to begin is by asking what is meant by the term "guidance." Much of the recent debate seems to have proceeded without a careful definition of the term, nor a well-spelled-out metric of how much of it is provided in a given instructional design. Reviewing the recent debate in *Educational Psychologist* (Kirschner et al., 2006; Hmelo-Silver et al., 2007; Kuhn, 2007; Schmidt, Loyens, van Gog, & Paas, 2007; Sweller, Kirschner, & Clark, 2007), the collection of instructional moves discussed under the rubric of guidance seems to include explanation, feedback, help, modeling, scaffolding, procedural direction, and others. Implicitly, guidance seems to have been taken as a superordinate category that describes *anything* an instructor provides to students to aid their learning and performance.

We find ourselves asking whether this is a useful concept to debate around—to us it seems problematic to meaningfully gauge "amount" across such a broad range of instructional moves. Should a high degree of scaffolding count as more or less guidance than a little explanation? Does a general model given ahead of time constitute more or less guidance than detailed feedback after the fact? Confusion over such questions may help to explain why the current debate has not reached resolution. Despite employing the common language of "guidance," the instructionist and constructivist camps appear to be using the term very differently. An indication that this may be the case can be seen in the different ways that each group approaches the fundamental problem of teaching complex knowledge and skills. Both groups recognize that novices have difficulty navigating large problem spaces (e.g., Sweller, 1988; Mayer, 2004; O'Neill & Weiler, 2006; Polman, 2000). However, the responses of constructivists and instructionists to this challenge are substantially different.

A key difference between instructionist and constructivist perspectives appears to center on the problem of part–whole relationships. In our interpretation, instructionists' main solution to the challenge of teaching a complex body of knowledge or skills is to break it down into smaller, clearer pieces, and provide up-front instruction about how to tackle each one. This way, students are not easily lost amid innumerable details. Eventually, of course, students have to put the pieces together to produce a whole, competent performance; but in the instructionist view, these part–whole relationships are best taught after each of the individual pieces has been mastered. Thus, from an instructionist perspective, a key purpose of instructional guidance is to reduce the extraneous cognitive load borne by the learner in processing each part of the material or task to be learned (Sweller, 1988).

From this perspective, it makes a lot of sense to focus on the amount of guidance provided in an instructional design. It also appears to be a straightforward matter. Having students figure out how to do something themselves entails the least guidance, and the greatest cognitive load. Scaffolding learners by providing a little support along the way entails moderate support, and leads to moderate cognitive load. Telling or showing students exactly how to accomplish a task provides maximum guidance, and introduces minimal cognitive load, which evidence shows can lead to certain kinds of learning gains (Kirschner et al., 2006).

We note, however, that what instructionists take as evidence of learning (at least in the experiments cited in the recent debate) is usually performance on a school-like task performed very shortly after initial training. For example, of the eight worked-example studies reviewed by Kirschner et al. (2006) that looked at transfer, all used transfer tasks that, at best, can be characterized as "very near" on most dimensions of Barnett and Ceci's (2002) transfer taxonomy. In all of these studies, the transfer tasks were identical to the learning tasks in terms of physical context (lab/classroom), modality (written task and problem format), and social context (individual). Most also used a very near temporal context (same session) with one experiment administering a transfer task the following day (Carroll, 1994, Exp. 1).

One important reason for constructivists' very different approach to teaching may come from their greater ambitions where transfer of learning is concerned. The common constructivist sentiment was expressed well by Barnett and Ceci (2005) when they wrote that "no one cares about learning if it stops at the schoolhouse door" (p. 295). This sentiment, so different from that embodied in instructionists' experimental protocols, has a large influence on constructivists' reading of the literature.

Alexander and Murphy (1999) stated that transfer of learning to novel contexts "has long been a problem of Gordian proportions for applied psychologists" (p. 561), and note that very little evidence of transfer to novel problems has been produced in the lab. However, Bransford and Schwartz (1999) have argued that many lab-based transfer studies make people "look dumb" (unable to transfer learning) because they exclusively evaluate sequestered problem solving. They describe sequestered problem solving as a type of testing in which learners are asked to directly apply learned material to a new context without any additional resources or the opportunity to learn in the new context by attempting and revising solutions (Bransford & Schwartz, 1999; Schwartz, Bransford, & Sears, 2005).

Bransford and Schwartz (1999) and Schwartz et al. (2005) describe a series of experiments demonstrating that more encouraging results are obtained when the concept of transfer is expanded beyond a direct application paradigm. They suggest thinking of transfer in terms of "preparation for future learning"—the ways in which a learning experience can prepare students to learn related ideas or skills more quickly or more deeply in the future. They further demonstrate that by measuring students' ability to flexibly respond to and learn from new situations after certain kinds of treatments, one can uncover positive transfer not detected by sequestered problem-solving tasks (Schwartz & Bransford, 1998; Schwartz & Martin, 2004).

These findings make a good deal of sense in relation to constructivist theories of memory, which suggest that better elaborated memories, with more extensive relational networks, will lead to more reliable recall of learned material (Bransford, 1979; Schank, 1982). These same theories suggest that learners will develop more elaborate memories, with more transfer-relevant relational networks, from more authentic experiences, particularly if the experiences are well-scaffolded.

To maximize the chances that worthwhile transfer will occur, constructivists argue that it is important to retain as much of the authenticity (and hence complexity) of the target task as is practical. This position is consistent with current theories of transfer, which suggest (for example) that "the primary job of transfer is the construction [by the learner] of an evolving representation of context that allows the knowledge user to make meaning of contexts" (Royer, Mestre, & Dufresne, 2005, p. xxii). The constructed understanding of a new context allows learners to bring prior knowledge to bear in an appropriate way.

This view of transfer helps to explain famously confusing findings such as those of Carraher, Carraher, and Schliemann (1985). They found that Brazilian children were capable of making sophisticated calculations in their heads when working as street vendors. However, the children could not successfully carry out the same calculations *with the very same quantities* when asked to do so in a lab situation where the problems were represented in writing, such as "5 × 35 = ?". When the same calculations were given in the form of word problems, the additional context helped the children to solve the problems with much greater success (Carraher et al., 1985). These findings illustrate the importance of seeing a context as similar to one previously encountered in order for transfer to occur.

Thus, while instructionists tend to view a complex problem-solving situation as a large problem space filled with extraneous demands on the learner's attention (e.g., Sweller, 1988), a constructivist is more likely to view it as a rich set of contextual cues that may later aid transfer (e.g., Brown, Collins, & Duguid, 1989). In keeping the "pieces" of the skill to be learned together in relation both to each other and a complex macro-context, constructivists aim to teach part–whole relationships throughout the learning process.

Jasper Woodbury is a good illustration of how differently constructivists think about the function of instructional materials. In this mathematics curriculum, students are provided with video case "anchors" that are intended to be authentic to real-life problems (Cognition and Technology Group at Vanderbilt, 1992). The math problems embedded in these video anchors are quite complex, and to an instructionist, likely appear utterly lacking in necessary instructional guidance. No doubt, by *itself* such a situation presents problems for learners, given the limitations of working memory (Cognition and Technology Group at Vanderbilt, 1994). But rather than being designed as a complete plan for instruction, the videos and associated materials were intended to create "teachable moments" in which guidance can be provided in a contextualized way that is responsive to the students' current frame of mind (Schwartz et al., 2005).

Saye and Brush (2002) have aptly described such responsive, real-time guidance from teachers as "soft scaffolding." As Pea (2004) explains, scaffolding is a kind of performance support that is adaptive to the learner's current capability, and gradually fades away. One of the main pedagogical functions of scaffolding is to channel or focus the learner's attention on what is important within a complex learning situation. While directing learner attention is a pedagogical function also endorsed by instructionists (Kirschner et al., 2006), constructivists seek to help learners identify important elements *in situ* rather than extracting them to present to learners.

This overview may help explain why it can appear to instructionists that constructivists are focused on minimizing guidance in instruction. We believe that this is not the case. For their part, constructivists are equally committed to helping novices cope with the limitations of working memory. However, rather than relying exclusively on shrinking the problem space and providing up-front instruction, they aim to help learners manage a large problem space by providing real-time supports that they theorize will make transfer to real-life situations more likely.

To be fair, constructivists have often not been as committed to measuring far transfer as they have been to advocating its importance. A thoroughgoing commitment to measuring unprompted far transfer would entail shadowing research participants in "real life," waiting for transfer to occur—an approach that to our knowledge has never been implemented, for obvious logistical reasons.

The Methodological Issue: Why "High" versus "Low" Guidance Studies Cannot Resolve the Debate

At first blush, judging the "success" of constructivist theory in informing instruction seems to imply a simple horse race: constructivism versus instructionism, winner takes all. In Western culture we seem drawn to contests between two extreme possibilities as a way to get at the truth of a matter, though this approach often does not produce the satisfaction we are after (Tannen, 1999). For example, while comparing test results from a constructivist classroom to those from a more traditional one tells us which class did better in those particular implementations, it does not generate the kind of evidence we need to make generalizable claims about causes. As Kirschner et al. (2006) point out, there are too many different variables at play, including the quality with which each approach is implemented, to make valid inferences about which factors are responsible for the differences. We acknowledge that a great deal of the evidence presented in support of constructivist teaching is subject to this critique. However, we argue

that the experimental "high versus low guidance" evidence introduced by scholars such as Kirschner et al. (2006) is also not a valid basis for making inferences about the fundamental merits of constructivist teaching.

Kirschner et al. (2006) use the level of guidance provided in instruction as a proxy by which to compare constructivist and instructionist approaches. The ensuing conversation about the efficacy of constructivist approaches to instruction has thus centered on the question of "how much" guidance is needed for effective instruction, and whether approaches such as guided inquiry and problem-based learning provide enough (Hmelo-Silver et al., 2007; Schmidt et al., 2007; Sweller et al., 2007). If one accepts the assumption that quantity of guidance is the most important difference between constructivist and instructionist designs, it then appears that *any* study comparing "low" and "high" guidance approaches is relevant to the discussion. However, as we have outlined above, the constructivist approach differs from the instructionist one not simply in the amount of guidance provided, but also in how problems are structured and the way in which guidance supports the learner. Importantly, the problem structure and the guidance provided are designed to be mutually supporting—theoretically, we would not expect one to be successful without the other.

Thus, while many of the studies that instructionists have cited in the recent debate are valuable for increasing our understanding about guidance within a particular framework (constructivist or instructionist), we submit that they are not valid evidence for making a judgment between the two frameworks. For example, Mayer (2004) reviewed a set of studies comparing pure-discovery learning with guided inquiry. However, discovery learning and guided inquiry are different forms of *constructivist* instruction. Neither falls under the rubric of explicit, up-front instruction. Thus at best, this study shows that within a constructivist framework, some scaffolding is better than none. It does not compare a constructivist approach with an instructionist one.

Similarly, the worked-example studies which Kirschner et al. (2006) review in detail (e.g., Sweller & Cooper, 1985; Cooper & Sweller, 1987) compare two variations on instructionist designs. One design involves showing students an example of how to do a problem, then gives them many opportunities to practice. The other design replaces a considerable portion of the practice problems with worked-out examples. In both cases, the instruction begins with showing a student exactly how to do a problem. Doing this with one example or with many constitutes a variation of the same instructionist approach. Thus, while useful in other ways, the results of these experiments also do not provide evidence on which to compare the relative success of constructivist and instructionist approaches.

The failure of both constructivists and instructionists to produce evidence that their counterparts find persuasive seems to be caused by a methodological catch-22. To wit: if one conducts a properly designed, classically controlled experiment varying *only* the amount of guidance provided in instruction, they are restricted to making comparisons *within* one of the two frameworks (instructionist or constructivist), or using a very impoverished version of one of the approaches. However, if we attempt to test a "good" instructionist lesson against a "good" constructivist one, we must involve differences in more than one variable, making our results ungeneralizable.

Reframing the Question

The methodological quandary discussed above is not new. The same issues were the focus of the 1987 debate between Papert and Pea around LOGO (Papert, 1987). Back then, Papert argued that it was meaningless to conduct a classical experimental trial of a learning innovation that, by its nature, requires a number of simultaneous and interdependent changes to instruction in order to be effective. The methodological debate remains very much alive today, mostly in the literature surrounding design-based research (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Collins, Joseph, & Bielaczyc, 2004); and there does not appear to be an end in sight.

Instructionists appear to be looking for a grand generalization: is constructivism or instructionism (as characterized by low or high guidance respectively) more effective? But perhaps this is not the most useful question to ask. Putting this in the context of the horse-race analogy, it is like asking which of two horses will win more races *on average*. But instructional designers and teachers deal in specifics. Their natural interest is in predicting what sort of instruction will be most effective in a *particular* situation.

If instructionists find this second question uncompelling, it may be because they believe they can develop a set of inviolate principles with which to predict the performance of instruction across all situations. However, there are a multitude of variables that affect the enactment of an instructional strategy (or a horse race) in a specific situation. For example, the experience of a teacher (Featherstone, 1992), the characteristics of the individual students (Jonassen & Grabowski, 1993), and the type of material to be learned (Stodolsky, 1988) all affect the implementation of instructional strategies in ways that can alter results. In both horse races and classrooms, we suggest that the art of picking a winner is not a matter of figuring out which "horse" has won the most in the past and betting on it in every race, but working to optimize one's chances of a positive outcome given what we know about the conditions in a particular situation.

How would the landscape of research change if instead of seeking rigid, universal *prescriptions*, we sought a tractable set of *considerations* for designing instructional guidance? With such an approach, instead of trying to develop absolute rules that apply to all instructional situations (e.g., more guidance is always better), we would develop principles that characterize successful guidance (e.g., the amount of guidance must be sufficient to support learners in seeing how the different pieces of the task fit together). The principles could then be used to consider a particular instructional context and learning goals, guiding design in a heuristic fashion. Instructionists have started us on this road by identifying one important consideration in instructional design—the amount of guidance provided. To this list, constructivists would add the context in which the guidance is given and the timing with which it is given.

In the following sections, we argue that the amount, context, and timing of instructional guidance are *all* important as factors to consider when planning instruction, regardless of whether one is coming from an instructionist or constructivist perspective. We do not aspire to provide (now or someday) a set of fully prescriptive rules about the "best" position along each dimension. Rather, as described above, we seek to develop principles for thinking about each factor that can be used to guide choices about guidance in specific learning situations.

Amount of Guidance: From Prescription to Guideline

The way the instructionist case has been argued to date, it would appear that more guidance is bound to be better in every case, regardless of the instructional goals being pursued or the domain in which they are being pursued (e.g., Sweller et al., 2007, p. 117). However, while we agree that amount of guidance is important to think about, we see the evidence as indicating that it is important in a different way. To us, the evidence suggests that the optimal amount of guidance often is an intermediate amount and the *granularity* of the advice provided in a design (i.e., the level of detail) is equally important.

To illustrate this point, we revisit one of the primary sources of evidence cited by Kirschner et al. (2006) for their claim that more guidance is always better, the worked-example effect. Of the nine studies they cite about the effect, five (Carroll, 1994; Cooper & Sweller, 1987; Paas & van Merriënboer, 1994; Sweller & Cooper, 1985; Trafton & Reiser, 1993) only compared *two* levels of guidance: fully worked-out examples and unsolved problems. While these studies can show the relative benefits of two particular "amounts" of guidance, they are not direct evidence of a continuous linear relationship between the two.

Looking at the studies cited that did have more than two conditions, additional guidance did not always result in a commensurate gain in learning (Miller, Lehman, & Koedinger, 1999; Paas, 1992; Quilici & Mayer, 1996). For example, Paas (1992) found that learners in a "completion" condition, who were given a problem only halfway worked out, performed as well on test problems as those given the problems fully worked out. In this case, the additional guidance given by the fully worked-out problem did not seem to provide additional learning benefits. Similarly, in a statistics problem-sorting task, Quilici and Mayer (1996) found that providing learners with three examples of each problem type, as opposed to one, did not result in any differences in students' ability to sort subsequent problems into the appropriate types.

This is not to say that additional guidance can *never* lead to learning gains—just that we cannot assume that it will *always* do so. In part, the learning gains achieved may depend on the purpose the additional guidance is serving. For example, in a worked-examples study asking learners to mentally translate between 2-D and 3-D representations of objects, Pillay (1994) found that worked examples showing three intermediate problem stages were more effective than those showing only one. She suggests this may be because in the three-stage representation, the distance between stages was small enough to "permit subjects to follow the transformation without having to generate [additional] stages themselves" (Pillay, 1994, p. 110). In a related finding, Catrambone (1994, 1995) found that when there are many individual steps involved, showing how they fit together to achieve sub-goals can make the steps more meaningful and useful. Finally, while not a worked-example study per se, the ninth study cited by Kirschner et al. (2006) looked at guidance in the form of a task for exploring a physics simulation microworld called *Electric Field Hockey* (Miller et al., 1999). They found that it was not the simple absence or presence of a guiding task, but specifically *how* the task influenced students' activities in the microworld that led to learning benefits.

Beyond the worked-example literature, there are other studies that suggest that aiming for the right level of granularity in guidance is a better guideline than "more is always better." For example, Nadolski, Kirschner, and van Merriënboer (2005) found that breaking the task of preparing a legal plea into four steps was more effective in supporting their target population of learners than presenting the task as a whole (one step) or in nine steps. While too few steps may leave learners unclear about what to do, too many seem to overwhelm them or prevent them from seeing the forest for the trees.

These findings are reminiscent of a much earlier study, in which Kittell (1957) compared three strategies for teaching learners to find the word that did not belong in a group, based on some underlying principle. In looking at multiple measures of retention and transfer, the most successful strategy was a medium-guidance one, which gave learners word sets and a general principle to apply. This outperformed both a low-guidance strategy that only gave learners the word sets, and a high-guidance strategy that gave learners the word sets, the general principle *and* an explanation of how the principle applied to each set. In this case, a very high level of guidance was unsuccessful—perhaps because it prompted learners to follow it rotely, without actively making sense of it (Mayer, 2004).

Altogether, this review of the evidence strongly suggests that while the amount of guidance in an instructional design is important to its effectiveness, more is not necessarily better in all cases.

Context of Guidance: Giving Students a Need to Know

As we discussed above, constructivists appear to pay more attention than instructionists to the context in which guidance is offered because they believe it to be important for retention and transfer. As support, constructivists cite studies indicating that when explanations are given in isolation, students learn very little, even when the explanations are well designed (e.g., Wieman & Perkins, 2005; Hrepic, Zollman, & Rebello, 2007). In higher education, for instance, it has been found that even just 15 minutes after a traditional lecture, students recall very little of the material covered (Wieman & Perkins, 2005). This can be the case even when students are given specific prompting about what to focus on (Hrepic et al., 2007). In Hake's (1998) large-scale study of physics learning across 62 universities, he found that students in traditional lecture classes recall an average of less than 30% of the concepts presented.

Several more detailed studies demonstrate how the context of an explanation affects learning results. Schwartz and Bransford (1998), for instance, demon-

strated that giving students a lecture on memory led to learning gains when students *first* had the opportunity to analyze results from a real memory experiment. However, when the distinctions between the cases were analyzed for them, or when the prior task involved summarizing a relevant text, the learning from the lecture was significantly less. Similarly, Capon and Kuhn (2004) found that giving students a relevant problem to solve before they heard a lecture on economics made them more likely to explain and apply the concepts on a later examination, instead of simply regurgitating the textbook definition.

We see a strong resonance between these findings and the constructivist perspective as we articulated it earlier. The implication seems to be that learners will construct a different understanding if they are given an explanation in isolation, versus first having an experience that gives them a "need to know." If students are provided with a meaningful goal, or at least the opportunity to develop one, it affects how they construct meaning from instruction, and thus what is learned (Miller et al., 1999; Schank, 1982).

A powerful illustration of the importance of context is provided by Schwartz and Martin (2004). In this study, ninth-grade students were taught statistics. After 2 weeks of instruction, half were asked to invent a method for comparing high scores from different distributions. The other half were shown how to solve the problem graphically and had an opportunity to practice. Students were further subdivided before being given a transfer problem in which they were asked to compute standardized scores numerically. Half the students in each condition were given a worked example for computing standardized scores, while the other half proceeded directly to the transfer test.

The best performers on the transfer post-test were those who were first asked to invent a method for comparing groups, and then given a worked example for computing standardized scores. This group scored twice as highly as each of the other groups, including the group that was given the tell-and-practice instruction followed by the worked example. As Schwartz et al. (2005) explain, the students who had a chance to invent a procedure first were able to learn more from the worked example.

These findings about the role of context in making guidance effective resonate in an interesting way with some of the findings from the worked-example literature. For example, in the original worked-example studies by Sweller and Cooper (1985) and Cooper and Sweller (1987), each worked example was given immediately prior to a similar problem that the student had to solve. While Sweller and Cooper (1985) describe this as a purely motivational strategy, Trafton and Reiser's (1993) work suggests that being able to access the worked example in the context of a problem to be solved is instrumental in achieving the learning benefits.

In the context of teaching LISP programming, Trafton and Reiser found that worked examples only conferred an advantage when they were presented immediately before a similar problem to be solved, and thus available in memory for use during practice (Trafton & Reiser, 1993). This suggests that at least part of the learning occurs when the students use the worked examples as aids in solving the subsequent problems. Further supporting the notion that worked examples provide a benefit as "on-line guides" during problem solving, Carroll (1994) observed that

> students in the worked example condition spent little time examining the worked example before attempting the accompanying practice problem ... instead they proceeded quickly to the practice problem and then referred back to the example as they wrote or completed the equation.
>
> (p. 364)

While the problem-solving context used in these studies is somewhat different than the kinds of context often used by constructivists (e.g., Pea, 1994; Schank & Neaman, 2001), they serve a similar purpose: to provide students with a goal or "need to know" that drives how the students construct an understanding from instruction. Thus, the worked-example literature, as well as the evidence presented by Schwartz and Martin (2004), suggest that the context in which worked examples are given strongly affects their efficacy.

Timing of Guidance: Is Sooner Always Better?

While not always in agreement about *when* guidance should be given, both constructivists and instructionists believe that the timing of instructional guidance is important (e.g., Anderson, 1993; Schwartz & Bransford, 1998). From an instructionist perspective, the best time to provide guidance is as soon as possible—either at the beginning of the instruction or as soon as a learner makes an error. However, the detailed research on intelligent-tutoring systems suggests that depending on the instructional goals being pursued, providing immediate guidance is not always the best strategy (Anderson, Corbett, Koedinger, & Pelletier, 1995; Anderson, 1993; Mathan & Koedinger, 2003).

Intelligent-tutoring systems are computer-based problem-solving environments that pose problems for learners, and offer individualized guidance based on observing every step of their attempts at solving it (Van Lehn, 1988). Intelligent advice from the computer is made possible by an "expert model" that is capable of solving all of the tutorial problems. The tutor thus "knows" whether or not each step taken by the learner is on a path to a valid solution or represents an error.

Early work by Anderson (1993) showed that offering learners guidance as soon as they strayed from a viable solution path increased their problem-solving speed. In a study using an intelligent tutor to teach students LISP programming, learners who were interrupted during their work and offered guidance as soon as they took a step that would not lead to a valid solution completed the programming exercises in about half the time taken by those who received feedback only on request (Anderson, 1993). Anderson et al. (1995) explained that the advantage of immediate feedback is that it is received when the learner's short-term memory retains enough of their attempted solution that they can understand and learn from the feedback.

On the other hand, in a more recent review of the intelligent-tutoring literature, Mathan and Koedinger (2003) suggest that delaying feedback may result in better retention and transfer of learning. As they explain, providing feedback as soon as an error is detected can rob learners of the opportunity to develop the evaluative skills needed to examine the effects of a problem-solving step, and attempt to repair it in case of error.

In an effort to test this hypothesis, Mathan and Koedinger (2003) conducted an experiment using two different feedback conditions. In one condition, the tutor offered immediate feedback on errors. In the other, the tutor waited to see whether learners detected their own errors, and attempted to guide them through detecting and correcting their mistakes *only* if they attempted to move on to a new problem. Findings from this study indicated that while the learners in the two groups performed similarly on the first problem, those in the delayed feedback condition learned at a faster rate on all subsequent problems (Mathan & Koedinger, 2003).

Together, the Anderson (1993) and Mathan and Koedinger (2003) studies suggest two things about the timing of instructional guidance. First, it clearly matters. When the type of feedback is held constant, timing strongly affects short-term outcomes. Second, the timing of feedback should vary according to instructional goals. While immediate feedback promotes more rapid problem solving in the short term, delaying feedback can result in better long-term retention and transfer.

The Possibility of a Joint Research Agenda

In the previous section we argued that amount, context, and timing are all important factors to attend to when planning instruction, whether one is coming from an instructionist or constructivist position. If we can talk about instructionist and constructivist instructional designs in similar terms, there may be potential for intellectual transfer between the groups.

Up till now, it seems that little cross-fertilization has occurred. Each group has been content to draw on and contribute to work by like-minded colleagues. For their part, instructionist scholars have expressed frustration that constructivists have "mostly ignored" the detailed findings of worked-example studies (Sweller et al., 2007, p. 119). Similarly, we are unaware of any major instructionist attempts to draw on constructivist findings. Despite this lack of crossover, the two camps often create guidance that is similar in its pedagogical functions. For example, both groups use some form of explanation (e.g., Garner, 1987; Schwartz & Bransford, 1998), procedural direction (e.g., van Merriënboer, 1997; McNeill, Lizotte, Krajcik, & Marx, 2006), and feedback (e.g., Phye & Sanders, 1994; Kolodner & Guzdial, 2000).

Looking at one example in particular, both Hmelo-Silver et al. (2007, p. 102) and Sweller et al. (2007, p. 118) note that instructionists' worked examples serve a similar function to real-time modeling in inquiry environments. Each kind of guidance shows learners how domain experts break down a problem to make it more soluble, and how the different concepts, strategies, and procedures they use are combined to achieve working solutions. Thus, perhaps the lack of crossover between the camps occurs not because there is no relevance, but because the findings of each are so ensconced in the language of their particular framework that these connections are difficult to make.

As a case in point, worked-example studies deal almost exclusively with well-structured problems in technical domains such as science and mathematics (see reviews by Sweller, van Merriënboer, & Paas, 1998; Atkinson, Derry, Renkl, & Wortham, 2000). The guidelines for the design of worked examples that have emerged from the research are thus worded specifically for the creation of algorithmic, written-out problem solutions. It is difficult for constructivists to transfer these findings to real-time guidance, and to the ill-defined problem domains (e.g., social studies, history, writing) with which many of them are concerned. Attempts to make such translations (e.g., Atkinson et al., 2000) run the risk of distorting the original approach. For example, Atkinson et al. (2000) describe how the guidelines for worked examples could be translated for use in the PBL environment of the Secondary Teacher Education Project (STEP). However, in many respects this description (for example, providing an expert solution before each case that students are given to solve) sounds more like up-front instruction than a problem-based approach.

We believe that the categories of amount, context, and timing may allow us to translate the worked-example findings in a more appropriate way and create an interesting and testable research agenda that could bring instructionists and constructivists together. We illustrate this in the following section. For instructionists, this agenda offers the opportunity to expand their work into new domains. For constructivists, it offers a strategy for drawing on a body of highly detailed findings that, as Kirschner et al. (2006) point out, have not been used much by constructivists in the past. Our example is just one of the many possible ways in which using the general language of amount, context, and timing may support productive conversations between instructionists and constructivists.

Applying Worked-Example Guidance Principles to Real-Time Modeling: Thinking about Amount, Context, and Timing

Earlier we noted that worked examples seem to be most effective when the *amount* of guidance they provide is sufficient to show learners how to get from one stage to the next (Pillay, 1994) and when individual steps are grouped together to show sub-goals along the way (Catrambone, 1994, 1995). In addition, it appears valuable to provide worked examples in the *context* of a problem to be solved (e.g., Trafton & Reiser, 1993) and perhaps *after* students have had a chance to attempt problems in the domain (Schwartz & Martin, 2004). What might such principles look like for real-time modeling? To give a concrete illustration, we consider the problem of teaching students to analyze the historical significance of a document. The first phase of the new research we envision could take an observational form. Outstanding history teachers could be identified (perhaps through large-scale test data) and observed over the course of several days. Observers would note the degree to which they use modeling to teach the analysis of historical documents, and whether the characteristics set forth in the worked-example literature are useful (and sufficient) to describe their practice.

This observational data would help us to describe more precisely how the principles described above are instantiated for modeling problem solving in complex and ill-defined domains. For example, when is the modeling provided, in what context is it provided, and how much of it is provided at one time?

The second phase of the research would use an experimental approach to build on the findings of the first. Systematic variations could be made in the amount, context, and timing of the modeling guidance. For example, in terms of timing, is it best for the teacher to model the analysis process (a) as a whole up-front; (b) as separate pieces before the students get to each part; (c) after students have had a chance to explore the problem; (d) only when learners commit an error (either immediately or after they have a chance to detect and correct their mistake); or (e) some combination of the above? In terms of the problem context, we could compare the following conditions: (a) simply providing a model; (b) providing the model in the context of another document to be analyzed; and (c) providing the model as guidance to help students use historical documents as evidence in a debate on a current governmental policy. For amount, we might look at variations in (a) the size of the sub-goals used; (b) how this structure is highlighted for students; and (c) the ways in which skilled teachers illustrate the transition from one stage of analysis to the next.

For each comparison outlined, research could examine both the learning process and learning outcomes. One outcome measure to examine would be students' ability to analyze other historical documents. Following Bransford and Schwartz's (1999) notion of transfer as preparation for future learning, this should be evaluated not only in terms of success in producing sophisticated solutions, but also how students go about attacking novel problems. For example, how do students adapt to analyze a *foreign* historical document? Are some groups more able to learn from initial mis-steps than others? In addition, we could examine the different ways in which students interacted with the teacher's modeling. Did they ask questions of the teacher? If so, what kind? If given the opportunity to access a video of the teacher's modeling, did they choose to use it? If so, when did they access it, what parts did they look at, and how did they try to use them?

Together, the kinds of studies we have outlined above may help to articulate principles for successful modeling of complex problem solving in ill-defined domains. Whether we can generate such principles more quickly by drawing on the worked-example findings is an interesting and potentially fruitful question. We would also expect such research to generate new avenues for investigation into the principles that make worked examples themselves effective. While there may be differences in what makes modeling and worked examples effective, we should be able to discuss them in similar terms. Such a research agenda could be of interest to instructionists and constructivists alike, and is just one example of how using a more general language can help to build common ground in the conversation on instructional design.

Conclusion

We have argued above that instructionists and constructivists may be talking past each other due to different ambitions with regard to transfer, and different assumptions about how best to make it happen. Instructionists seem to be interested primarily in transfer to formally similar problems in school-like situations, while constructivists are more interested in cross-contextual transfer and transfer between school-like and real-world situations. These very different commitments may have led them to different readings of the literature, different demands on theory and research methods, and different approaches to supporting novices in complex problem spaces.

In our view, the multiple differences in the way guidance is conceived by each group has led to a methodological stalemate, in which each group brings forth evidence that the other group does not find convincing. We suggest that it is time to move away from the unproductive adversarial question of "which approach is better overall" to a collaborative agenda in which we seek to understand principles for designing the amount, context, and timing of guidance that is suitable for particular goals and situations. To illustrate what this might mean, we have sketched the outlines of a research agenda that we believe may bring instructionists and constructivists together in a productive partnership.

We note that some important factors in the design and effectiveness of instructional guidance have been overlooked in this chapter, including how the principles outlined may interact with individual differences between learners (Jonassen & Grabowski, 1993). We expect that as research progresses, these will be elucidated. We also note that proceeding down the path we have outlined leaves a greater degree of inference to the designer in translating principles to practice. We feel that this flexibility is necessary to design instruction that meets the needs of particular learning situations; however, we wonder how willing instructionists will be to slacken their commitment to generating deterministic prescriptions for instruction.

We hope that this discussion has provided a helpful way to think about the current debate surrounding instructional guidance. As we move forward in the 21st century, educational researchers are recognizing the inherent complexity involved in many of the most important educational issues. Answers seem to lie not in contests between two extreme possibilities, but in the fine distinctions that emerge from a detailed and multidimensional examination of the evidence. This framework presents one tool to help researchers make progress on this long-term endeavor.

Question: Clark. *I enjoyed reading your chapter. We all hope that a clearer conception of instructional guidance will be the main product of the debate. In the 1970s Lee Cronbach complained that we measure individual differences in micrometers and instructional methods with divining rods. The situation is more or less the same today. You argue persuasively for systematic research that clarifies issues such as the amount, context, and timing of guidance. Yet it also seems that there is no agreement about the cognitive functions of the various operations that are offered to define guidance in research or practice. How can we accomplish what you propose without considering the function of the enormous range of instructional support activities that are considered to be examples of "guidance" by various groups in education?*

Reply: Wise and O'Neill. We agree completely that a better theorization (and testing) of the functions that different types of guidance play in supporting the learning process is needed. In fact, in an earlier version of our chapter we used pedagogical function as a grouping category to show commonalities in the different kinds of guidance favored by the two camps. To us, this seems like a very productive avenue for starting to make sense of the large range of instructional support activities currently employed and studied. Furthermore, once function is used to organize the various types of guidance, it appears that much of the debate is not about *whether* a particular type of guidance (pedagogical function) should be given but *how* it should be given. For example, we discussed how both worked examples and real-time modeling in inquiry environments serve the function of showing learners how domain experts break down a problem to make it more soluble and how the different concepts, strategies, and procedures they use are combined to achieve working solutions. To us, it appears that the difference between worked examples and real-time modeling is not the pedagogical function served, but whether the guidance is given before or in the course of problem solving and in the set up of the problem-solving situation itself (i.e., the timing and context of the guidance).

Similarly, both instructionists and constructivists employ some form of guidance to provide procedural direction to students, either in the form of process worksheets or as scaffolds to support the inquiry process. Again, the pedagogical function in both cases seems the same: to provide students with steps, rules, or guidelines to follow in completing a task; the difference lies in the level of granularity with which the guidance is given and the larger task structure in which the guidance is embedded (i.e., the amount and context of the guidance).

We do not assume that the most effective way to give guidance will be the same across types of guidance (or across learning goals for a particular kind of guidance). Thus, identifying a pedagogical function and the various forms of guidance that seem to be fulfilling this function is an important prerequisite to conducting this kind of research. Additionally, in conducting the research on how students interact with the guidance, we can examine if the guidance is indeed serving the pedagogical function we anticipated, or if there are other function(s) being served.

Question: Clark. *Isn't it possible that farther transfer is possible when instructionist-guided support employs what has been called varied (or variable) practice and haven't most of the reviews of past research on this issue concluded that the evidence only supported gradual and limited transfer despite claims to the contrary?*

Reply: Wise and O'Neill. The reviews of research on this topic, including the ones we cover in our chapter, do recommend caution in our expectations of transfer. As Anderson, Reder, and Simon (1996) point out, depending on the relationship between the learning and transfer materials (and, we would add, learning and transfer situations) "there can be either large amounts of transfer, a modest amount, no transfer at all or even negative transfer" (pp. 7–8). As we read the reviews, however, they do not argue that only small, gradual transfer is possible. On the contrary, Anderson et al. (1996) present evidence to *contest* the claim that knowledge does not transfer between tasks. Similarly, Perkins and Salomon (1989) sum up their review by stating that "Given appropriate conditions, such as cueing, practicing, generating abstract rules, socially developing explanations and principles, conjuring up analogies ... and the like, transfer from one problem domain to another can be obtained" (p. 22).

We do take issue with one small point in Perkins and Salomon's summation: the idea that what is transferred are "general skills." In our view, there is not good evidence for the existence of cognitive skills that are truly general from the time that they are mastered, though there is great evidence that such skills can be *generalized* over time with appropriate kinds of practice, in the appropriate contexts, and with appropriate scaffolding and feedback. To borrow Perkins and Salomon's example, the chess master *can* potentially learn to transfer his skill in chess to battlefield strategy; but this will not necessarily be a quick process, however straightforward the analogies may seem to an outside observer.

As you suggest, part of the training that may help to generalize skill is variable practice—though we don't see this as an instructional technique that instructionists have a monopoly on. However, from a constructivist perspective the practice should include not only variations in the cognitive features of the task, but also variation across a range of appropriate situations of use. So, the bookish chess master may need considerable practice not only in working a range of military strategy problems, but working them in a variety of appropriate social settings (such as a hotly contested argument in a battlefield tent with a group of gruff generals) before he can reliably transfer his strategic thinking to be productive in a military context.

Question: Rosenshine. *I was very impressed with the examples of instructional procedures related to the amount of guidance, the context of guidance, and the timing of guidance. The example of giving students a problem to solve before they hear a lecture is a lovely and inventive idea. But every one of the examples in these three sections came from a study in which the criterion was "gain in learning." I see absolutely no conflict between constructivists and direct-instruction advocates when we read, discuss, or conduct research in which "gain in learning" is the criterion. Do you agree that when constructivists and direct-instruction advocates are both interested in the outcome of a gain in learning there is no longer any conflict between the two groups?*

Also, in the opening section, you note that constructivists are interested in "worthwhile transfer," "preparation for future learning," the "ability to learn from new situations," and "transfer to real-life situations." But, you also note that "constructivists have not been as committed to measuring far transfer as they have been to advocating its importance." If constructivists can't measure far transfer and worthwhile transfer, why do you have a section on this topic? Wouldn't it be more honest to drop these unsupported claims and focus only on constructivist instructional procedures that lead to gains in learning? Would your argument suffer if inferences about "worthwhile transfer" were removed from your chapter?

Reply: Wise and O'Neill. In our minds there is a difference in what instructionists and constructivists are talking about when they refer to a gain in learning. As we wrote in our chapter, it seems to us that constructivists are more ambitious with respect to transfer, and link their understanding of "learning gain" very closely with it. In this vein, we aim to optimize instructional activities for depth of understanding and far transfer rather than quick learning and near transfer. We feel that there are two very compelling reasons why we need to include a consideration of far transfer when we talk about gains in learning: first, this is the ultimate goal of education—to prepare students to use the knowledge they learn in school in real-world contexts at some point in the future; second, instructional techniques that lead to the quickest initial learning may not necessarily be the same as those that lead to the best retention and long-term transfer. For example, we discuss the intelligent-tutoring literature in which immediate feedback has been found to produce more rapid learning initially, while delayed feedback has led to better retention and transfer. Thus, tasks that test students on exactly what they learned or provide a weak measure of near transfer cannot be used as a proxy for far-transfer results, and more far-transfer measures need to be included in future studies.

It seems that you over-interpret the statement we made in our chapter about the difficulties constructivists have had in assessing far transfer. We certainly did not intend to say that far transfer is unmeasurable. Nor did we intend to say that constructivists have entirely failed to measure far transfer. Our comments in this section were meant to point out the logistical, conceptual, and monetary challenges of measuring the extreme end of the transfer continuum—unprompted use of knowledge in everyday settings, long after the initial learning. Barnett and Ceci (2002) have devised a useful taxonomy for thinking about the different ways a transfer situation can differ from the original learning situation in terms of six dimensions: the knowledge domain; physical context; temporal context; functional context; social context; and task modality. In the great majority of the instructionist studies we have reviewed, "transfer" is operationalized as testing performance in a context that differs on only one of these dimensions, knowledge domain, and only slightly at that. The other five dimensions are virtually the same in the training and testing scenarios: physical context (lab/classroom); functional context (academic, evaluative); social context (individual); temporal context (same day or next day); and modality (written/typed task, same problem format). Even within the knowledge domain, what instructionists seem to consider "far" transfer is quite different from what constructivists do.

While we think that both instructionists and constructivists could and should go farther in terms of measuring transfer along all of Barnett and Ceci's dimensions, there are some notable examples of far-transfer measurements that have been done. In science, researchers have looked for evidence of transfer from project-based learning in the quality of their written reports, and their ability to plan novel investigations (O'Neill & Polman, 2004). Klahr and Nigam (2004) used a similar strategy when they asked students who had been learning a control-of-variables strategy to critique other students' (problematic) science fair posters. Fong, Krantz, and Nisbett (1986) conducted one of the few studies that looked at transfer outside the classroom, by calling research participants at home and conducting a "survey" that asked them questions that involved the statistical knowledge they had been learning in school. Studies like these demonstrate that it is neither impossible nor impractical to devise and implement far-transfer measures in educational research.

Question: Rosenshine. *I enjoyed reading your example about learning how to teach students to analyze historical documents. You wrote that the first step would be to find outstanding history teachers and observe their instructional procedures. The second step would be experimental studies where students were taught to analyze documents using ideas learned from observing these teachers and, within the intervention, there might be variations in the timing and use of models. These two steps seemed very similar to direct instruction and process–product research where the researchers also identified outstanding teachers, studied their instructional procedures, and then used these procedures in experimental studies where a measure of student achievement was the criterion. In other words, if I were interested in training students to analyze historical documents I would want to use the same procedures that you describe. So if the goal is to train students to become more proficient at analyzing historical documents, do you see any conflict, or any difference, in how constructivism and direct-instruction advocates would try to achieve this goal?*

Reply: Wise and O'Neill. We thank you for pointing out the (unintentional) similarities between our proposed new line of research on modeling in history instruction and the procedures of process–product research of 30 years ago. In addition to the important similarities you point out, we think that what we are proposing in our chapter is different in some key ways from that earlier work. Perhaps it would be fair to say that our proposal adopts characteristics of the successful process–product research design, while adapting them to address current research problems. For example, the process–product research appears to be largely domain general and to have examined the presence or absence of a wide array of instructional procedures. In contrast, the work we propose is domain-specific, focuses on one particular form of instructional guidance (modeling), and examines the enactment (the "how") rather than the simple presence or absence. These are important differences in our view since the past three decades of cognitive science have shown how powerful domain-specific influences are on thinking and learning. Further, the series of experiments we are suggesting appear (from our understanding) to address questions of a more specific nature than the process–product research.

Finally, we believe some important changes have taken place in history instruction which would likely have substantial influence on both the implementation and effectiveness of modeling in history classrooms. The original process–product research seems to have compared various forms of teacher-centered classrooms. Until the late 1980s or early 1990s, relatively little attention was paid in North America to teaching historical interpretation; emphasis on the mastery of textbook narratives seems to have prevailed. That situation has now changed—the 1996 revision of the US National Standards for History included substantial recommendations for conveying and understanding the nature of historical knowledge. Today, regular publications for teachers frequently contain articles on teaching history using primary-source evidence.

To sum up, we want to highlight that our goal in suggesting this program of research was not to create a "constructivist" approach to teaching students how to analyze historical documents, but to suggest a joint research agenda that both constructivists and instructionists could pursue. The goal would be to catalog the many ways in which modeling guidance is given in history classrooms, with the aim of optimizing the use of this powerful instructional strategy for different instructional contexts and goals.

Question: Rosenshine. *When students construct knowledge on their own, or with minimal supervision, they are also likely to also construct errors. How do constructivists handle this possibility of students making unintentional but persistent errors under conditions of minimal supervision?*

Reply: Wise and O'Neill. Given the constructivist belief that students inevitably construct their own understandings (regardless of what form of instruction is used), it is certainly possible that at some points they will construct problematic understandings. We agree that this is an important issue. Your question seems to imply, however, that this is a special problem for constructivist classrooms because they leave students alone with instructional materials and activities for too long. While students in a constructivist classroom may spend a great deal of their time working on a problem with their classmates, this does not mean that they are without teacher supervision. As we pointed out in our chapter, constructivists are not advocating "minimal guidance," we are advocating that guidance be more situated, flexible, and responsive.

We expect three features of a well-designed constructivist classroom to safeguard against students holding on to misinterpretations that arise during instruction. First, a well-designed constructivist activity will give students opportunities to test the viability of their understandings as they develop. For example, in the case of science, the instructional design may include specific experimentation that targets common student preconceptions. Second, a constructivist classroom that promotes valued forms of student discussion, justification, and testing of ideas (scientific argumentation), will provide additional opportunities for problematic ideas to surface and be transformed. Third, in both cases, a skilled teacher circulating in the classroom will be able to provide just-in-time guidance (in the moment it is needed) to help students understand their errors or misunderstandings, and correct them.

References

Alexander, P. A., & Murphy, P. K. (1999). Nurturing the seeds of transfer: A domain-specific perspective. *International Journal of Educational Research, 31*, 561–576.

Anderson, J., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167–207.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher, 25*(4), 5–11.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*, 181–214.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.

Barnett, S. M., & Ceci, S. J. (2005). Reframing the evaluation of education: Assessing whether learning transfers beyond the classroom. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 295–312). Greenwich, CT: Information Age Publishing.

Bransford, J. D. (1979). *Human cognition: Learning, understanding, and remembering.* Belmont, CA: Wadsworth.

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

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction, 22*, 61–79.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in the schools. *British Journal of Developmental Psychology, 3*, 21–29.

Carroll, W. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 86*, 360–367.

Catrambone, R. (1994). Improving examples to improve transfer to novel problems. *Memory and Cognition, 22*, 606–615.

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87*, 5–17.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cognition and Technology Group at Vanderbilt. (1992). The Jasper experiment: An exploration of issues in learning and instructional design. *Educational Technology Research & Development, 40*(1), 65–80.

Cognition and Technology Group at Vanderbilt. (1994). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1), 15–42.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347–362.

Featherstone, H. (1992). *Learning from the first years of classroom teaching: The journey in, the journey out* (NCRTL Special Report). East Lansing, MI: National Center for Research on Teacher Learning, Michigan State University.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18*, 253–292.

Garner, R. (1987). Strategies for reading and studying expository text. *Educational Psychologist, 22*(3–4), 299–312.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74.

Hickey, D. T., Kindfield, A. C. H., Horwitz, P., & Christie, M. (1999). Advancing educational theory by enhancing practice in a technology supported genetics learning environment. *Journal of Education, 181*(2), 25–55.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99–107.

Hrepic, Z., Zollman, D. A., & Rebello, N. S. (2007). Comparing students' and experts' understanding of the content of a lecture. *Journal of Science Education and Technology, 16*(3), 213–224.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Kittell, J. E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. *Journal of Educational Psychology, 48*(7), 391–405.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*, 661–667.

Kolodner, J. L., & Guzdial, M. (2000). Theory and practice of case-based learning aids. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 215–242). Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, D. (2007). Is direct instruction an answer to the right question? *Educational Psychologist, 42*(2), 109–113.

Mathan, S., & Koedinger, K. R. (2003). Recasting the feedback debate: Benefits of tutoring error detection and correction skills. In U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (pp. 13–20). Amsterdam: IOS Press.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14–19.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153–191.

Miller, C., Lehman, J., & Koedinger, K. (1999). Goals and learning in microworlds. *Cognitive Science, 23*, 305–336.

Nadolski, R. J., Kirschner, P. A., & van Merriënboer, J. J. G. (2005). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology, 75*, 223–237.

O'Neill, D. K., & Polman, J. L. (2004). Why educate "little scientists"? Examining the potential of practice-based scientific literacy. *Journal of Research in Science Teaching, 41*(3), 234–266.

O'Neill, D. K., & Weiler, M. J. (2006). Cognitive tools for understanding history: What more do we need? *Journal of Educational Computing Research, 35*(2), 179–195.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434.

Paas, F., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*, 122–133.

Papert, S. (1987). Computer criticism vs. technocentric thinking. *Educational Researcher, 16*(5), 22–30.

Pea, R. D. (1994). Seeing what we build together: Distributed multimedia learning environments for transformative communications. *The Journal of the Learning Sciences, 3*(3), 285–299.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related concepts for learning, education and human activity. *The Journal of the Learning Sciences, 13*(3), 423–451.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18*(1), 16–25.

Phye, G. D., & Sanders, C. E. (1994). Advice and feedback: Elements of practice for problem solving. *Contemporary Educational Psychology, 19*(3), 286–301.

Pillay, H. (1994). Cognitive load and mental rotation: Structuring orthographic projection for learning and problem solving. *Instructional Science, 22*, 91–113.

Polman, J. L. (2000). *Designing project-based science: Connecting learners through guided inquiry.* New York: Teachers College Press.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*, 144–161.

Royer, J. M., Mestre, J. P., & Dufresne, R. J. (2005). Introduction: Framing the transfer problem. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. vii–xxvi). Greenwich, CT: Information Age Publishing.

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research & Development, 50*(3), 77–96.

Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people.* Cambridge: Cambridge University.

Schank, R. C., & Neaman, A. (2001). Motivation and failure in educational simulation design. In K. D. Forbus & P. J. Feltovich (Eds.), *Smart machines in education: The coming revolution in educational technology* (pp. 37–69). Cambridge, MA: MIT Press.

Schmidt, H. G., Loyens, S. M. M., van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 91–97.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*, 475–522.

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). Greenwich, CT: Information Age Publishing.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition and Instruction, 22*, 129–184.

Stodolsky, S. S. (1988). *The subject matters: Classroom activity in math and social studies.* Chicago, IL: University of Chicago Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*, 59–89.

Sweller, J., Kirschner, P. A., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist, 42*(2), 115–121.

Sweller, J., van Merriënboer, J. G., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Tannen, D. (1999). *The argument culture: Stopping America's war of words.* New York: Ballantine Books.

Trafton, J. G., & Reiser, B. J. (1993). *The contribution of studying examples and solving problems to skill acquisition.* Paper presented at the 15th Annual Conference of the Cognitive Science Society.

Van Lehn, K. (1988). Student modeling. In M. C. Polson & J. J. Richardson (Eds.), *Foundations of intelligent tutoring systems* (pp. 55–78). Hillsdale, NJ: Erlbaum.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills.* Englewood Cliffs, NJ: Educational Technology Publications.

Wieman, C., & Perkins, K. (2005). Transforming physics education. *Physics Today, 59*(11), 36–41.