

ANALYZING CLAIMS ABOUT COGNITIVE DEMAND AND STUDENT LEARNING

Samuel Otten

University of Missouri
ottensa@missouri.edu

Zandra de Araujo

University of Missouri
dearaujoz@missouri.edu

Corey Webel

University of Missouri
webelcm@missouri.edu

A commonly accepted claim in mathematics education is that there is a relationship between the cognitive demand of mathematical task enactments and students' subsequent learning. One study often cited to support this claim is Stein and Lane (1996), and in 44% of those citations, Stein and Lane (1996) is the sole reference provided. Citation analysis reveals that many of these claims go beyond the warrants provided by the Stein and Lane study, either by granting more confidence in the relationship than the study design allows or by phrasing the claim causally. A few other studies are occasionally cited in conjunction with Stein and Lane (1996), but there remains a need for replication studies to provide better empirical support for claims about cognitive demand and student learning and to refine our shared understanding.

Keywords: Cognition, Instructional Activities and Practices, Research Methods

Replication studies are rarer in education research than in other fields, leading Makel and Plucker (2014) to call for more replications because such studies can both identify and remedy methodological biases and can be instrumental in buttressing robust findings or clarifying inconsistent findings. In these ways, replications can play a role in the field's systematic accumulation of knowledge (National Research Council, 2002). An area ripe for replication would be a testable, widely-held belief that is resting on a relatively inadequate empirical foundation. One belief seemingly shared by most scholars in mathematics education is that the cognitive demand of mathematical tasks and task enactments (Stein, Grover, & Henningsen, 1996) is important with respect to student learning outcomes. It may be that experiences with cognitively demanding tasks lead to positive learning outcomes or it may be that having students experience cognitively demanding tasks is an end itself. The latter is a philosophical position based on values, whereas the former is a testable position that currently rests on some supporting evidence, but what is the extent of that evidence? In our own past work related to cognitive demand (e.g., de Araujo, 2012; Otten & Soria, 2014), we noticed an extensive literature base on the nature of cognitive demand (Doyle, 1983; Stein, Smith, Henningsen, & Silver, 2009) and factors influencing cognitive demand throughout mathematical task implementations (Jackson, Garrison, Wilson, Gibbons, & Shahan, 2013; Wilhelm, 2014), but a weaker empirical foundation for the direct link between cognitive demand and student learning.

The hypothesis motivating this study was that a single reference—Stein and Lane (1996)—constituted a large portion of the warrants for claims in the mathematics education literature about the link between cognitive demand and learning. If this hypothesis were true, then it would become imperative to critically analyze the research design, evidence, and claims made in Stein and Lane (1996) and to consider possibilities of replication. We examined the claims for which Stein and Lane (1996) was included as a citation and we identified other references, if they existed, that were also cited for those same claims. In the following sections, we briefly summarize the Stein and Lane (1996) study, describe our method for compiling and analyzing citations to Stein and Lane (1996), and then share our key results.

Summary of Stein and Lane (1996)

Stein and Lane (1996) stems from a project well known in mathematics education—Quantitative Understanding: Amplifying Student Achievement and Reasoning (QUASAR). QUASAR involved a

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

university partnership with six urban middle schools with the overall goals of promoting reform-oriented mathematics instruction and investigating the feasibility of such instruction in schools with a history of poor mathematics performance (see Silver & Stein, 1996, for an overview). Within that context, Stein and Lane (1996) sought to “present evidence regarding the degree to which the presence of reform features of instruction are linked to increases in student understanding of mathematics” (p. 51). Their study focused on 4 of the 6 middle schools from the QUASAR project over a three-year period. Data consisted of narrative summary field notes and video recordings of three three-day observation cycles in three teachers’ classrooms in each school each year, a classroom observation instrument completed based on the field notes and video recordings, and Fall and Spring administrations of a project-developed assessment instrument (Lane, 1993). Mathematical tasks were identified and of the 620 main tasks, a stratified random sample of 144 was drawn and the task set-up and task implementation of these 144 tasks were coded for cognitive demand. Levels of cognitive demand were collapsed to 2 (high and low). A 25% sample of the 144 tasks was double coded with 79% agreement.

The assessment instrument consisted of 36 open-ended tasks distributed into four forms (9 questions per form) and a 5-point scoring rubric (0–4) for each task. Analysis focused on 11 of the tasks and used not the scores themselves but “the average percentage of student responses across tasks that were scored at the two most proficient score levels (3 or 4)” (p. 68) and how this average percentage shifted between Fall Year 1 and Spring Year 3.

To generate their findings, Stein and Lane rank ordered the four schools based on their gains in percentage of students at the top two levels of proficiency and then focused on Site A, which had gained the most (36%), and Site D, which had gained the least (17%). They compared these learning gain rankings with the school profile for task enactments and noted the following:

Site D’s profile can be seen as embodying a more conventional mathematics program in which many or most tasks lent themselves to being solved with a single strategy, using only one representation (usually symbolic), and without much explanation and/or discussion. Site A’s profile, on the other hand, suggests a well-functioning reform program that is successfully utilizing tasks that invite and support students’ use of multiple solution strategies and multiple representations, along with student discussion of their work. (p. 71)

In other words, the tasks in Site D classrooms were often set-up and implemented at low levels of cognitive demand whereas tasks in Site A classrooms were often set-up and implemented at high levels, and Site D had the lowest proficiency gains whereas Site A had the highest.

The fact that the sites conformed to a positive relationship between cognitive demand and gain scores seems, on the surface, to be compelling. The number of sites, however, is quite low for making even correlational claims about the relationship between task features and student achievement. Moreover, the use of only 11 items to measure learning over a three-year period gives pause. In addition, sites were only compared in relation to each other, not to a standardized score, raising questions about the size of the differences between the outcomes of the different sites. The study also did not account for differences among teachers within the same site. The authors acknowledged “the possibility that the findings of this study may partially reflect differences in school-level variables in addition to the documented differences in instructional practices” (p. 75), but in other instances the authors shifted from discussing the results in terms of correlations to suggesting causation. For example, the authors drew the following conclusions:

[S]tudents appear to **benefit more from** inconsistently implemented tasks that began with the encouragement to use multiple solution strategies, multiple connected representations, and explanations, than they do from tasks that—from the start—required only single solution

strategies, single representations, and little or no mathematical communication (p. 74, emphasis added)

The authors appropriately hedged this statement with the word “appear” but the language in bold suggests not just that Site A used more high demand tasks and performed better, but that the use of high demand tasks was a primary reason for higher student achievement. This claim seems to extend beyond the actual warrants of the study given its limitations. Yet, as we show below, the hedged claims appear to have been accepted widely by several researchers and national organizations alike, frequently in an unhedged fashion. Although Stein and Lane recommended that a replication of their study be carried out, to our knowledge this has not occurred except a substantially modified replication (Otten, 2012) involving 12 middle school classrooms in which cognitive demand was decidedly not correlated with measured student learning.

We wish to emphasize that by pointing out some of the limitations of the study, we do not mean to criticize the study itself but rather to raise caution about the claims that can be made—by Stein, Lane, or others—based on this single study’s evidence. These cautions led us to investigate the claims that have been made based on Stein and Lane’s (1996) work.

Method

Compiling Citations of Stein and Lane (1996)

Similar to Leatham and Winiecke (2014), we used the “cited by” tool on Google Scholar to obtain a list of articles in which Stein and Lane (1996) was cited. This initial search yielded 298 sources. Because we were interested in how the field of mathematics education specifically draws upon this reference, we retained only those articles that were published in mathematics education journals that received an A-grade in Toerner and Arzarello (2012, December) and also appeared as a top-five journal in the rankings compiled by Nivens and Otten (in press). These journals were *Educational Studies in Mathematics*, *Journal for Research in Mathematics Education*, *Mathematical Thinking and Learning*, *Journal of Mathematical Behavior*, *Journal of Mathematics Teacher Education*, and *ZDM – The International Journal of Mathematics Education*. This constraint yielded 26 articles. We then expanded our search beyond the Google Scholar results to include other forms of codified mathematics education literature—namely, the *Second Handbook of Research on Mathematics Teaching and Learning* (Lester, 2007) and two major policy documents from NCTM (2000, 2014). This additional search led to the identification of 4 handbook chapters and one policy document (NCTM, 2014) that cited Stein and Lane (1996). The 31 analyzed resources are marked with * in the reference list.

Analyzing Citations of Stein and Lane (1996)

The goal of our analysis was to understand the claims for which authors cited the Stein and Lane (1996) study. Within the 31 sources, we located the Stein and Lane (1996) citations in text, yielding 60 excerpts. We then used analytic memos to briefly describe the claim that was being supported by the Stein and Lane citation and met to identify the recurring themes and further refine them into codes. We came to a consensus on four codes (see Table 1), which we used to code all but one of the excerpts.

To gain a more nuanced understanding of the ways in which authors used the Stein and Lane reference, we subdivided the learning claim excerpts into two groups based on whether the claim was causal or non-causal. For example, the following excerpt was coded causal because the authors’ state that tasks result in an increase in student understanding:

As the research conducted by the QUASAR project indicates, when teachers choose tasks that require a high-level of cognitive demand, set them up and implement them in ways that maintain

a high-level of cognitive demand, **the result is** an increase in student understanding and reasoning (Stein & Lane, 1996). (Arbaugh & Brown, 2005, p. 527, emphasis added)

In the same article, we coded an earlier reference as non-causal:

The relationship between the types of tasks students engage in when learning mathematics and the mathematics they learn has been a subject of research for many years (see, for example, Hiebert & Wearne, 1993; Marx & Walsh, 1988; Stein & Lane, 1996). These research studies indicate that **a relationship exists** between the level of student thinking required by a mathematical task and the nature of students' understanding of mathematics. (Arbaugh & Brown, 2005, p. 505, emphasis added)

Table 1: Coding Scheme for Citations of Stein and Lane (1996)

| Code Descriptions | Examples |
|--|---|
| Claim about a relationship between tasks and student learning [learning claim] | “There is evidence that solving a task of high cognitive demand or a cognitively demanding task (CDT) has a positive impact on students’ conceptual understanding (Stein & Lane, 1996).” (Wilhelm, 2014, p. 637) |
| Claim that tasks (or task implementations) and learning have been studied [study claim] | “Research has focused on instructor questions across the K–16 spectrum and examined ... the value of questions to student learning (Silver, 1996; Stein & Lane, 1996).” (Fukawa-Connelly, 2012, p. 332) |
| Claim about the levels of cognitive demand (but no connection to student learning) [cognitive demand claim] | “[Stein and Lane] use four categories: Memorization, Procedures without Connections, Procedures with Connections, and Doing Mathematics.” (White & Mesa, 2014, p. 678) |
| Claim about a research method used [method claim] | “Of particular note are the procedures developed by Stein and Lane (1996) and Stein et al. (1996) for sampling and coding mathematical tasks and linking those findings to student outcomes.” (Gearhart et al., 1999, p. 309) |

We also distinguished four levels of attribution within the learning claims, from weak to strong—(1) the author(s) state that others have made claims about the relationship but the authors do not necessarily endorse the claims themselves, (2) the authors state the relationship with an explicit hedge (e.g., it is “suggested” or there is “possibly” a relationship) or as being found specifically in the context of the cited study, (3) the relationship is stated as something found in past studies and the authors explicitly or implicitly endorse the findings beyond the cited study’s context, and (4) the relationship is stated as a generalized fact. We independently assigned each of the learning claim excerpts a level of attribution and met to discuss any differences until we came to a consensus.

Results

Table 2 contains frequencies for the codes described in Table 1. The most common claim for which Stein and Lane (1996) was used as support was the notion that the cognitive demand of mathematical tasks or task implementations is somehow related to student learning. More than three-quarters (77.4%) of the resources made claims of this sort, encompassing 60% of the total Stein and Lane citations analyzed. Of the 36 learning-claim citations of Stein and Lane (1996), 16 (44.4%) cited only Stein and Lane. This supports our hypothesis that Stein and Lane (1996) often stands alone as the empirical basis for such claims. A substantial minority (25%) of the claims (spanning 41.9% of

the resources) did not involve student learning but instead focused solely on the construct of cognitive demand or task implementation. Six claims (10%) simply acknowledged that studies such as Stein and Lane (1996) existed and two (3.3%) referenced the type of study design that Stein and Lane employed. Overall, the citations of Stein and Lane (1996) predominantly involved the relationship between cognitive demand and student learning.

Table 2: Frequencies of Claims for which Stein and Lane (1996) was Cited

| Code | Number of Excerpts | Number of Sources |
|-------------------------|--------------------|-------------------|
| Learning claims | 36 (60.0%) | 24 (77.4%) |
| Study claims | 6 (10.0%) | 6 (19.4%) |
| Cognitive demand claims | 15 (25.0%) | 13 (41.9%) |
| Method claims | 2 (3.3%) | 1 (3.2%) |
| Other | 1 (1.7%) | 1 (3.2%) |

Of the 36 learning claims, 25 (69.4%) were correlational claims and 11 (30.6%) were causal. Although correlational claims are more justified than causal claims, few alluded to limitations of the Stein and Lane (1996) study. Furthermore, we found 11 citations that made causal claims. For example, Wilhelm (2014) stated, “There is evidence that solving a task of high cognitive demand or a cognitively demanding task (CDT) **has a positive impact** on students’ conceptual understanding (Stein & Lane, 1996)” (p. 639, emphasis added).

Returning to the 36 learning claims overall, Table 3 summarizes their attribution level. More than half (58.3%) of the claims were Level 2—that is, rephrasings of Stein and Lane (1996) or hedged claims that stated what Stein and Lane (and possibly others) had found, without implying that it was a generalized result. For example, Otten and Soria (2014) wrote that “**Stein and Lane (1996) argued** that maintaining high cognitive demand has positive benefits with respect to student learning” (p. 816, emphasis added). Such instances are defensible because it is true that Stein and Lane argued for the benefits of high cognitive demand tasks.

Table 3: Attribution Levels for Learning Claims Supported by Stein and Lane (1996)

| Attribution Level | Number of Excerpts | Number of Sources |
|--------------------------------|--------------------|-------------------|
| 1 (non-endorsed) | 4 (11.1%) | 4 (12.9%) |
| 2 (hedged or context-specific) | 17 (47.2%) | 10 (32.3%) |
| 3 (study-supported) | 9 (25.0%) | 8 (25.8%) |
| 4 (general fact) | 6 (16.7%) | 5 (16.1%) |

Yet, 41.7% of the learning claims went further by referring in some sense to a general relationship (level 3 or 4 attribution). For example, NCTM (2014) wrote about the relationship in a general fashion, as a matter of fact:

Learning is greatest in classrooms where the tasks consistently encourage high-level student thinking and reasoning and least in classrooms where the tasks are routinely procedural in nature (Boaler & Staples, 2008; Hiebert & Wearne, 1993; Stein & Lane, 1996). (p. 17)

In this case, NCTM did not mention cognitive demand specifically but instead mentioned key features related to cognitive demand. They also cited studies in addition to Stein and Lane (1996) as support for their claim, but due to space limitations, we will not describe those studies here.

Discussion

Our citation analysis revealed that, across primary journals in our field, Stein and Lane (1996) was used to support a substantial number of claims about the link between cognitive demand of mathematical tasks and students' mathematical learning. Of those claims, 55.6% also involved additional references beyond Stein and Lane, but these were often policy or practitioner works rather than empirical research, perhaps aligning with a philosophical stance on cognitive demand but not an empirical one. The U.S. Department of Education and the National Science Foundation (2013) described six types of education research: foundational, exploratory, design and development, efficacy, effectiveness, and scale-up. Exploratory research "examines relationships among important constructs in education and learning to establish logical connections that may form the basis for future interventions or strategies to improve education outcomes" (p. 9). They indicate that exploratory "connections are usually correlational rather than causal" (p. 9). Certainly, the work of Stein and Lane (1996) and others (e.g., Hiebert and Wearne, 1993; Tarr et al., 2008) provides valuable information about a potential relationship between cognitive demand and student learning, but based on the evidence, we judge this work to be at no higher than the exploratory level. Replications are needed to test whether this relationship holds under various conditions (design and development research), and a significant amount of work would need to be done in order to make causal claims that could estimate the average impact of using high cognitive demand tasks. As Stein and Lane said in 1996, "the analyses discussed herein should be replicated" (p. 75), but rather than answering this call, the field has instead justified the belief that cognitively demanding tasks relate with (or cause) higher student learning outcomes by drawing on studies in the initial stages of the research progression.

Though we argue for the need of later-stage research studies to examine the link between cognitively demanding tasks and student learning, we do not dismiss the value of smaller-scale studies. We recognize the importance of a wide array of research; however, certain types of claims (e.g., factors that generally relate to measurable student outcomes) are best supported empirically by larger-scale study designs. If we include such claims in our work, we should be cognizant of the nature of the empirical support, and modify claims accordingly. And if claims are value-based rather than empirical, that is appropriate, but they should be written as such.

At the center of this article is a deep question about the relationship between cognitively demanding mathematical tasks and students' mathematical learning, broadly construed. Although we have critiqued some of the sources of evidence for claims made about this relationship, it is not our intention to cast doubt on the relationship itself. In fact, we believe it is highly likely that a positive relationship does exist. Yet, as a field, we should not be satisfied with shared beliefs based on insufficient evidence. Instead, we should strive for a body of evidence that would convince not only someone who is already predisposed to value cognitively demanding experiences but that would convince a skeptic. Therefore, we join Makel and Plucker (2014) and Warne (2014) in encouraging editors and agencies to open the door to replications, even—or perhaps especially—if it investigates claims that many already take to be true.

References (* indicates sources included in the analysis)

- *Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education*, 8, 499–536.
- *Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, 40, 119–156.
- *Boston, M. D., & Smith, M. S. (2011). A 'task-centric approach' to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks. *ZDM – The International Journal on Mathematics Education*, 43, 965–977.

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

- *Cai, J. (2005). US and Chinese teachers' constructing, knowing, and evaluating representations to teach mathematics. *Mathematical Thinking and Learning*, 7, 135–169.
- de Araujo, Z. (2012). *Transferring demand: Secondary teachers' selection and enactment of mathematics tasks for English language learners*. Unpublished doctoral dissertation. University of Georgia. Athens, GA.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53, 159–199.
- *Drollinger-Vetter, B., Lipowsky, F., Pauli, C., Reusser, K., & Klieme, E. (2006). Cognitive level in problem segments and theory segments. *ZDM – The International Journal on Mathematics Education*, 38, 399–412.
- *Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Charlotte, NC: Information Age.
- *Fukawa-Connelly, T. P. (2012). A case study of one instructor's lecture-based teaching of proof in abstract algebra: making sense of her pedagogical moves. *Educational Studies in Mathematics*, 81, 325–345.
- *Gearhart, M., Saxe, G. B., Seltzer, M., Schlackman, J., Ching, C. C., Nasir, N., . . . Sloan, T. F. (1999). Opportunities to learn fractions in elementary mathematics classrooms. *Journal for Research in Mathematics Education*, 30, 286–315.
- *Grover, B. W., & Connor, J. (2000). Characteristics of the college geometry course for preservice secondary teachers. *Journal of Mathematics Teacher Education*, 3, 47–67.
- *Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.
- Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, 30, 393–425.
- *Hsu, H. Y., & Silver, E. A. (2014). Cognitive complexity of mathematics instructional tasks in a Taiwanese classroom: An examination of task sources. *Journal for Research in Mathematics Education*, 45, 460–496.
- *Huang, R., & Cai, J. (2011). Pedagogical representations to teach linear relations in Chinese and US classrooms: Parallel or hierarchical? *Journal of Mathematical Behavior*, 30, 149–165.
- Institute of Education Sciences, & National Science Foundation. (2013). Common guidelines for education research and development Retrieved March 25, 2016, from <http://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf>
- *Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education*, 44, 646–682.
- Lane, S. (1993). The conceptual framework for the development of a mathematics performance assessment. *Educational Measurement: Issues and Practice*, 12, 16–23.
- *Leatham, K. R., Peterson, B. E., Stockero, S. L., & Van Zoest, L. R. (2015). Conceptualizing mathematically significant pedagogical opportunities to build on student thinking. *Journal for Research in Mathematics Education*, 46, 88–124.
- Leatham, K. R., & Winiecke, T. (2014). The case of the Case of Benny: Elucidating the influence of a landmark study in mathematics education. *Journal of Mathematical Behavior*, 35, 101–109.
- Lester, F. K., Jr. (Ed.). (2007). *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43, 304–316.
- *Mamolo, A., Ruttenberg-Rozen, R., & Whiteley, W. (2015). Developing a network of and for geometric reasoning. *ZDM*, 47, 1–14.
- *McCaffrey, L. S., Hamilton, B. M., Stecher, S. P., Klein, D., & Bugliari, A. R. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32, 493–517.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- *National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematics success for all*. Reston, VA: Author.
- National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- *Nie, B., Freedman, T., Hwang, S., Wang, N., Moyer, J. C., & Cai, J. (2013). An investigation of teachers' intentions and reflections about using Standards-based and traditional textbooks. *ZDM*, 45, 699–711.

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

- Nivens, R. A., & Otten, S. (in press). Assessing journal quality in mathematics education. *Journal for Research in Mathematics Education*.
- Otten, S. (2012). Exploring cognitive demand, student participation, and learning in middle school algebra. In L. R. Van Zoest, J.-J. Lo & J. L. Kratky (Eds.), *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 359–362). Kalamazoo, MI: Western Michigan University.
- *Otten, S., & Soria, V. M. (2014). Relationships between students' learning and their participation during enactment of middle school algebra tasks. *ZDM*, *46*, 815–827.
- *Ramful, A. (2014). Reversible reasoning in fractional situations: Theorems-in-action and constraints. *Journal of Mathematical Behavior*, *33*, 119–130.
- *Remillard, J. T., Harris, B., & Agodini, R. (2014). The influence of curriculum material design on opportunities for student learning. *ZDM*, *46*, 735–749.
- *Silver, E. A. (2009). Cross-national comparisons of mathematics curriculum materials. *ZDM*, *41*, 827–832.
- *Silver, E. A., & Herbst, P. G. (2007). Theory in mathematics education scholarship. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 39–67). Charlotte, NC: Information Age.
- Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education*, *30*, 476–521.
- *Steele, M. D., & Rogers, K. C. (2012). Relationships between mathematical knowledge for teaching and teaching practice: the case of proof. *Journal of Mathematics Teacher Education*, *15*, 159–180.
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, *33*, 455–488.
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, *2*, 50–80.
- *Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Charlotte, NC: Information Age.
- Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional development* (2nd ed.). New York, NY: Teachers College Press.
- *Stockero, S. L., & Van Zoest, L. R. (2013). Characterizing pivotal teaching moments in beginning mathematics teachers' practice. *Journal of Mathematics Teacher Education*, *16*, 125–147.
- Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, *39*, 247–280.
- *Thanheiser, E. (2014). Developing prospective teachers' conceptions with well-designed tasks: explaining successes and analyzing conceptual difficulties. *Journal of Mathematics Teacher Education*, *18*, 141–172.
- Toerner, G., & Arzarello, F. (2012, December). Grading mathematics education research journals. *Newsletter of the European Mathematical Society*, *86*, 52–54.
- *Ubuz, B., Erbaş, A. K., Çetinkaya, B., & Özgeldi, M. (2010). Exploring the quality of the mathematical tasks in the new Turkish elementary school mathematics curriculum guidebook: the case of algebra. *ZDM*, *42*, 483–491.
- Warne, R. T. (2014). Two additional suggested reforms to encourage replication studies in educational research. *Educational Researcher*, *43*, 465.
- *Warshauer, H. K. (2014). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, *17*, 1–26.
- *White, N., & Mesa, V. (2014). Describing cognitive orientation of Calculus 1 tasks across different types of coursework. *ZDM*, *46*, 675–690.
- *Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education*, *45*, 636–674.