



Improving the accuracy of teachers' judgments of student learning

Keith W. Thiede^{*}, Jonathan L. Brendefur, Michele B. Carney, Joe Champion, Lindsey Turner, Roger Stewart, Richard D. Osguthorpe

Boise State University, USA

HIGHLIGHTS

- Professional development in student-centered instruction improved the accuracy of teachers' judgments of student learning.
- Professional development on formative assessment did not affect the accuracy of teachers' judgments.
- Accuracy of teacher judgments were associated with student achievement.
- Professional development in student-centered instruction of mathematics also improved student achievement.
- Relative accuracy of teacher judgments mediated the effect of professional development on student achievement.

ARTICLE INFO

Article history:

Received 20 April 2018
Received in revised form
20 June 2018
Accepted 16 August 2018

ABSTRACT

We examined the effect of different professional development programs on the accuracy of teachers' judgments of their students' learning during three academic years. Teachers participated in a program focused on improving (a) use of formative assessment, (b) student-centered mathematics instruction, (c) use of both formative assessment and student-centered mathematics instruction; or (d) neither—a control group. Teachers' judgment accuracy was greater for teachers who participated in professional development around improving student-centered mathematics instruction than for other groups. A multilevel analysis showed a significant positive relation between the accuracy of teachers' judgments and student achievement.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Models of instruction describe teaching as the interplay between teachers' judgments of student learning and subsequent instruction (e.g., Box, Skoog, & Dabbs, 2015; Donovan, Bransford, & Pellegrino, 2000; Ready & Wright, 2011; Ruiz-Primo & Furtak, 2007). That is, teachers' judgments of student learning guide instruction (e.g., Alvidrez & Weinstein, 1999). For instance, judgments of student learning help teachers identify struggling students for additional instruction (e.g., Bailey & Drummond, 2006), and shaping feedback to students (e.g., Hoge & Coladarci, 1989). Given the role of teacher judgments in instruction, it is important to find ways to improve the accuracy of teachers' judgments. The present investigation evaluates the efficacy of different professional development programs on teachers' judgment accuracy.

^{*} Corresponding author. College of Education, Boise State University, 1910 University Drive, Boise, ID 83725, USA.

E-mail address: keiththiede@boisestate.edu (K.W. Thiede).

Research on teacher judgments has been conducted in a number of countries. Südkamp, Kaiser, and Möller (2012) noted in their recent review of the teacher judgment literature that a substantial amount of research has focused on how the accuracy of teacher judgments is affected by the characteristics of the judgment, test, students, and teachers (see also Hoge & Coladarci, 1989). This research, for instance, has shown that judgment accuracy is better when teachers are informed about the nature of the test prior to judging student performance than when they are uninformed (e.g., Feinberg & Shapiro, 2009) and when tests are more closely related to the classroom curriculum than when they are more general standardized tests (e.g., Jenkins & Demaray, 2016). It has also shown that judgment accuracy is better when teachers judge performance of higher achieving students than lower achieving students (e.g., Hurwitz, Elliot, & Braden, 2007), and judgment accuracy is not influenced by teaching experience (Zhu & Urhahne, 2015). However, relatively little research has examined ways to improve judgment accuracy—which is the focus of the present study. Before reviewing the literature focused on improving judgment accuracy, we first describe how judgment accuracy has been measured and

operationalized in the literature.

1.1. Measuring judgment accuracy

The standard procedure for gathering teacher judgments involves showing teachers a test and asking them to predict how their students will do on the test. Students then complete the test. Judgment accuracy is operationalized as the match between predicted and actual performance.

Judgment accuracy that describes the degree to which the magnitude of judgments is related to the actual magnitude of students' performance on a test is called *absolute accuracy* (Dunlosky & Rawson, 2012) and is often reported as confidence bias (i.e., average predicted performance minus average actual performance computed across students). Teachers typically over estimate students' performance, which is reflected in positive confidence bias (e.g., Rausch, Karing, Dörfler, & Artelt, 2016). For instance, over half the teachers in Bates and Nettelbeck (2001) were over-confident by 9–12 months on standardize reading assessment. *Relative accuracy* describes the degree to which judgments discriminate between different levels of performance across students. This is typically reported as the intra-teacher correlation between predicted and actual performance computed across students (e.g., Helmke & Schrader, 1987). Relative accuracy varies dramatically across studies. Accuracy reported in the meta-analysis conducted by Südkamp et al. (2012) varied from $r = 0.80$ (Methe, Hintze, & Floyd, 2008) to $r = -0.03$ (Graney, 2008).

It is important to note that absolute accuracy and relative accuracy are statistically independent of one another (Dunlosky & Thiede, 2013). To illustrate this point, we have created data for four teachers, each of whom is predicting the performance of five students on a test with 10 questions. As seen in Table 1, Teacher 1 is neither over-confident nor under-confident because the confidence bias is zero; this represents perfect absolute accuracy. Teacher 1 also has perfect relative accuracy—as the correlation between predicted and actual performance computed across the five students is $+1.0$. Teacher 2 is over-confident by an average of five points, but the correlation between predicted and actual performance is $+1.0$; thus, Teacher 2 has poor absolute accuracy, but perfect relative accuracy. Teacher 3 has perfect absolute accuracy (as the average predicted performance minus the average actual performance equals zero), but has relative accuracy of -1.0 (which is perfectly inaccurate). Teacher 4 has poor absolute accuracy and poor relative accuracy.

To illustrate the link between judgment accuracy and instruction, consider instructional decisions for these four teacher. Regarding absolute accuracy, Teachers 2 and 4 are quite over-confident. They think their students will get, on average, seven of ten points on the test, which may be well enough for the teacher to

proceed to the next topic. However, this overconfidence may lead the teachers to advance before their students have adequately learned the materials; therefore, foundational ideas may remain unlearned. By contrast, Teachers 1 and 3 may pace instruction more appropriately because their judgments accurately reflect their students' performance. Regarding relative accuracy, Teachers 3 and 4 did not accurately discriminate the students who have learned more from those who have learned less. These teachers might allocate more time to work individually with Students 4 and 5 (who actually did better on the test) rather than work with Students 1 and 2 (who actually did worse on the test). By contrast, Teachers 1 and 2 had perfect relative accuracy. The students they judged to have better learned the materials in fact did better on the test and the students they judged to have less mastery of the materials in fact did worse on the test; therefore, they are more likely to work with students who would benefit from additional instruction. Thus, judgment accuracy can influence decisions about pacing of instruction (absolute accuracy) and decisions about tailoring instruction to the needs of individual students (relative accuracy) (Thiede, Oswalt, Brendefur, Carney, & Osguthorpe, in press).

An alternative to asking teachers to predict performance on the test as a whole is to have teachers predict performance item by item for each student (for a discussion see Coladarsi, 1986). This approach makes it possible to compute accuracy as a hit rate across individual items. A teacher with a high hit rate can accurately judge how a student will do on each item. Knowing how each student would perform on each item would allow a teacher to tailor instruction to the item level.

1.2. Improving judgment accuracy

The cue-utilization framework (Koriat, 1997) provides a model for understanding how to improve monitoring accuracy. According to the framework, monitoring accuracy is driven by the cues people use to make judgments. Monitoring accuracy should improve when people have access to cues that are diagnostic of subsequent test performance (e.g., Thiede, Griffin, Wiley, & Anderson, 2010; Hertzog, Dunlosky, Robinson, & Kidder, 2003) and then use these cues to make judgments (Koriat & Bjork, 2005; Van Loon, de Bruin, van Gog, van Merriënboer, & Dunlosky, 2014). The present investigation examined how judgment accuracy was affected by two approaches used to increase access and use of diagnostic cues.

Use of formative assessment. (This section is now before the section on student-centered instruction.) Past performance is often the best statistical predictor of future performance. Teacher judgments are influenced by students' past performance (Hecht & Greenfield, 2002; Martínez, Stecher, & Borko, 2009). That is, teachers may use scholastic history in general (Dusek & Joseph, 1983) or performance on formative assessments conducted

Table 1
Illustrations of the relation between absolute and relative accuracy.

Performance	Teacher 1		Teacher 2		Teacher 3		Teacher 4	
	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual
Student 1	9	9	9	4	9	5	9	0
Student 2	8	8	8	3	8	6	8	1
Student 3	7	7	7	2	7	7	7	2
Student 4	6	6	6	1	6	8	6	3
Student 5	5	5	5	0	5	9	5	4
Average	7	7	7	2	7	7	7	2
Accuracy								
Absolute	0		+5		0		+5	
Relative	+1.0		+1.0		-1.0		-1.0	

Download English Version:

<https://daneshyari.com/en/article/10127048>

Download Persian Version:

<https://daneshyari.com/article/10127048>

[Daneshyari.com](https://daneshyari.com)