education policy analysis archives

A peer-reviewed, independent, open access, multilingual journal

epaa aape

Arizona State University

Volume 24 Number 86 August 8, 2016 ISSN 1068-2341

Science Teachers' Data Use Practices: A Descriptive Analysis

Virginia Snodgrass Rangel
University of Houston
&
Carlos Monroy
Rice University
&
Elizabeth R. Bell
Accelerate Learning Inc.
United States

Citation: Rangel, V. S., Monroy, C., & Bell, E. (2016). Science teachers' data use practices: A descriptive analysis. *Education Policy Analysis Archives, 24*(86). http://dx.doi.org/10.14507/epaa.24.2348

Abstract: There is a debate on students' low science achievement in the United States, particularly among low income, African American students, and Latino students. An important part of the education community's response to low achievement generally and in science specifically has been the implementation of high stakes accountability policies. Because of accountability's emphasis on educator data use, much research has examined different facets of it throughout educational organizations, but research has not focused on the extent to which data use might be content-specific. The purpose of this paper, then, was to investigate the data use practices of science teachers. Drawing from a broader study of science teachers in grades 5-8 across six school districts, this study reports results from teacher surveys and interviews. The findings indicate that while there

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 16/11/2015
Revisions received: 21/4/2016
Accepted: 21/4/2016

were examples of science-specific data use, most of the science teachers used data in ways consistent with previous content-agnostic research. Implications for future research, policy, and practice are discussed.
Keywords: Data use; science education; assessment; accountability

Las practicas del uso de data por profesores de las ciencias: Un análisis descriptiva
Resumo: Hay un debate sobre el rendimiento bajo en las ciencias de estudiantes en los Estados Unidos, especialmente entre los de bajos ingresos, afro-americanos y latinos. Un parte importante de la respuesta por parte de la comunidad educativa a la cuestión del rendimiento bajo, (en general y en las ciencias en particular), ha sido la implementación de políticas de altos riesgos de responsabilidad. Debido al énfasis de la responsabilidad en el uso de los datos por el educador, muchas investigaciones han examinado facetas diferentes de eso a lo largo de organizaciones educativas, pero las investigaciones no han enfocado en la medida en que el uso de datos podría ser específico al contenido. El propósito de este artículo, entonces, fue investigar las prácticas de uso de datos por profesores de las ciencias. Derivado de un estudio más amplio de profesores de las ciencias de grados 5-8 de en seis distritos escolares, este artículo se informe de los resultados de encuestas y entrevistas con profesores. Los resultados indican que mientras que hay ejemplos de usos de datos específicos a las ciencias, la mayoría de los profesores de ciencias usó los datos en maneras agnósticas del contenido, es decir, en maneras consistentes con investigaciones anteriores. Se discuten las implicaciones para investigaciones, políticas y prácticas en el futuro.
Palabras claves: Uso de datos; enseñanza de las ciencias; evaluación; responsabilidad

Práticas de uso de dados de professores de ciências: Uma análise descritiva
Resumo: Há um debate sobre o baixo desempenho dos estudantes em ciências nos Estados Unidos, particularmente entre alunos de baixa renda, afro-americanos e latinos. Uma parte importante da resposta da comunidade educativa para o baixo desempenho geral e em ciências especificamente tem sido a implementação de políticas de responsabilidades de altos riscos. Por causa da ênfase de responsabilidade em uso de dados do educador, muitas pesquisas examinaram diferentes facetas disto através de organizações educacionais, mas a pesquisa não focou na medica em que o uso de dados pode ser de conteúdo específico. O propósito deste artigo, então, era de investigar as práticas de uso de dados dos professores de ciências. A partir de um estudo mais amplo de professores de ciências na 5 à 8 série em seis distritos escolares, este estudo relata resultados de pesquisas e entrevistas de professores. Os resultados indicam que, enquanto há exemplos de uso de dados específicos em ciências, a maioria dos professores de ciências usaram dados de maneiras consistentes com pesquisas anteriores de conteúdo agnóstico. Implicações para futuras pesquisas, políticas e praticas são discutidos.
Palavras-chave: Uso de dados; educação científica; avaliação; responsabilidade

Introduction

There is a debate on students' low science achievement in the United States, particularly among low income students, African American students, and Latino students. Though the average eighth grader performed slightly better on the NAEP science test in 2011 than in 2009, African American and Latino students as well as students eligible for the federal free or reduced lunch program scored 27-35 points below White students (NCES, 2011). An important part of the education community's response to low achievement generally and in science specifically has been the implementation of high stakes accountability policies. Implemented at the state level beginning in the 1980s, accountability policies, in theory, should lead to school improvement by focusing educators' attention on student outcomes, as measured by standardized assessments (O'Day & Smith, 1993; Smith & O'Day, 1990). Implied by these policies is that educators will use the data from the standardized assessments to improve how schools and districts meet their students' needs (Hamilton & Stecher, 2006; Hamilton, Stecher, & Klein, 2002; Hanushek & Raymond, 2005). The No Child Left Behind Act (NCLB), passed in 2001, established accountability as federal policy as well.

Because of accountability's emphasis on educator data use, much research has examined different facets of it throughout educational organizations. Researchers have described the kinds of data teachers use as well as the varied ways in which teachers use those data (Beaver & Weinbaum, 2015; Datnow, Park, & Wohlstetter, 2007; Gallagher, Means, & Padilla, 2008; Hamilton, Berends, & Stecher, 2005; Lachat & Smith, 2005; Marsh, Pane, & Hamilton, 2006; Wayman & Stringfield, 2006; Wayman, Cho, Jimerson, & Spikes, 2012; Wayman, Snodgrass Rangel, Jimerson, & Cho, 2010). They have identified several factors that facilitate data use, such as data management systems (Gallagher et al., 2008; Lachat & Smith, 2005; Long, Rivas, Light, & Mandinach, 2008; Means et al., 2010; Supovitz & Klein, 2003; Wayman, Stringfield, & Yakimowski, 2004), structured time for reflection (Wayman et al., 2010), opportunities to learn how to use and interpret data (Wayman et al., 2010; Jimerson & Wayman, 2015), and supportive district (Coburn, Touré, & Yamashita, 2009; Wayman, Cho, & Shaw, 2009; Young, 2006) and school leadership (Copland, 2003; Lachat & Smith, 2005; McLaughlin & Mitra, 2003; Wayman et al., 2010; Young, 2006).

They also have found other factors that inhibit teacher data use, such as a lack of supportive school or district leadership (Wayman et al., 2012), a lack of alignment between school and district leaders around data use (Wayman et al., 2010; Young, 2006), problems with technology (Breiter & Light, 2006; Cho & Wayman, 2014; Coburn, 2005; Gallagher et al., 2008; Marsh et al., 2006; Means et al., 2010; Wayman, Cho, & Johnston, 2007; Wayman et al., 2012), such as a lack of integration across data systems, and the perceived usefulness or trustworthiness of data (Breiter & Light, 2006; Ingram, Seashore Louis, & Schroeder, 2004; Kerr et al., 2006; Marsh et al., 2006; Valli & Buese, 2007; Young, 2006). Other work has investigated the factors that shape teachers' data use and decision making, including teachers' beliefs (Coburn, 2001; Coburn & Turner, 2011; Palmer & Snodgrass Rangel, 2011; Spillane, Reiser, & Reimer, 2002) and communities (Daly, 2012; Huguet, Marsh, & Farrell, 2014; Marsh, Bertrand, & Fuguet, 2015; Nelson, Slavit, Perkins, & Hathorn, 2008), instructional coaches (Marsh, McCombs, & Martorell, 2010; Marsh et al., 2015; Swinnerton, 2007; Young, 2006) and principals (Ingram et al., 2004; Wayman & Stringfield, 2006; Wayman et al., 2007; Wayman et al., 2009), district context and leadership (Coburn et al., 2009), and the policy context (Datnow et al., 2012; Horn, Kane, & Wilson, 2015).

Most research on educator data use, however, either has been agnostic to the particular subject area from which the data are derived or has focused on data use in English language arts or mathematics. As a result, what scholars have reported as the ways in which teachers use data also has focused on 'general' data use or on the use of reading and/or mathematics data. Given that until 2007 accountability policies such as NCLB only required annual testing in English language arts and mathematics, this narrower view makes sense. Beginning in 2007, however, NCLB required states to test students in science once in elementary school, once in middle school, and once in high school. Furthermore, the unique nature of science (e.g., the heavy emphasis on scientific practices in addition to scientific knowledge) warrants additional investigation into the specific data use practices of science educators.

This study also was carried out in a rapidly changing data use landscape, namely because of the rapid expansion of technology in classrooms, including more sophisticated data systems and digital curricula. The mandate for more and better educator data use led to an expansion in the use of data warehouses (Kerr et al., 2006; Wayman et al., 2004; Wayman & Stringfield, 2006) and, more recently, in the creation of program-based student management systems[1], such as the one embedded in the online science curriculum that was the focus of the broader study from which these findings are drawn.

Therefore, the purpose of this paper is to investigate the data use practices of science teachers. Working from the premise that science is unique when compared to other disciplines because of the focus on hands-on, inquiry-based learning (Minner, Levy, & Century, 2010) and performance-based assessments (NRC, 2000; Songer & Ruiz-Primo, 2012), this exploratory study set out to understand what data science teachers use, in what ways, and what data they want to use. This study focused on science teachers in grades 5-8 for two reasons. First, beginning in sixth grade and sometimes even as early as fifth grade, science teachers usually are departmentalized and only teach science and so they teach science on a regular basis, unlike many elementary school teachers (Milner, Sondergeld, Demir, Johnson, & Czerniak, 2012). Second, the middle school years in particular are a crucial time for students in science because research suggests that it is then that students cement either their interest (Auger & Blackhurst, 2005; Bandura, Barbaranelli, Caprara, & Pastorelli, 2001; Riegle-Crumb, Moore, & Ramos-Wada, 2011; Tai, Liu, Maltese, & Fan, 2006) or disinterest (Aschbacher, Li, & Roth, 2010; NRC, 2011) in science. In order to guide the investigation, we used the following research questions:

1. What data do science teachers use in their classrooms?
2. What data do science teachers want to use in their classrooms?

In the remainder of this section, we review the literature on data use and accountability, and then teacher data use more specifically. Then we describe some of the 'best practices' in science instruction and assessment and explain why these might lead to unique data use practices in science classrooms, and what those practices might look like.

Literature Review

Test-based Accountability and Data Use

Accountability rests on the premise that the right policies can induce teachers and educators to make *better* instructional decisions based on data. As a result, major state and federal education reform efforts have included the creation of accountability systems that utilize high stakes standardized tests in certain subject areas to focus educators' attention on student learning outcomes (Hanushek, 2003; Hanushek & Raymond, 2002; O'Day & Smith, 1993; Opfer, Henry, & Mashburn, 2008). Educational data use is a key tenet of these accountability policies. In theory, educator analysis of the student achievement data collected each year is meant to improve teacher practice, student learning, and how schools function generally (Hamilton & Stecher, 2006; Hamilton et al., 2002; Hanushek & Raymond, 2005; O'Day & Smith, 1993; Smith & O'Day, 1990). The reauthorization of ESEA in 2001 (NCLB), in particular, placed renewed emphasis on using student

---

[1] By program-specific data management systems, we are referring to data systems that are embedded in specific curricula reather than general systems that are program and content agnostic, such as Schoology.

data to improve educational decision-making and student outcomes (Fuhrman & Elmore, 2004; Lachat & Smith, 2005; Supovitz, 2009; Wayman, 2005).

Teacher Data Use

While there has been little research on the data use of science teachers specifically, much research has explored teacher data use generally.

Kinds of data teachers use. Research consistently has found that teachers use a wide variety of student data in their work, though most of this research has not clarified whether the data teachers use are specific to particular subject areas. These data fall into two general categories: Achievement or performance data and non-academic data. Within the general performance data category, teachers' data use breaks down even further depending on how the data are collected. By far the most common source of data for teachers is state achievement tests (Beaver & Weinbaum, 2015; Gallagher et al., 2008; Marsh et al., 2006; Wayman, Cho, Jimerson, & Snodgrass Rangel, 2015), and from district or industry-created interim progress or benchmark tests (Marsh et al., 2006), which often are designed to prepare students to take end-of-year state assessments (Supovitz, 2012). Indeed, in their recent study of state performance data, Beaver and Weinbaum (2015) described their use by teachers as "pervasive".

Teacher academic data use is not limited to state or interim assessments, however, but also includes classroom work and observations of students during class. Classroom work includes teacher- or campus-made tests and quizzes as well as homework (Gallagher et al., 2008; Marsh et al., 2006; Wayman et al., 2010). Other classroom data that teachers have reported collecting and using include observations of student work in class and student questioning and discussions (Marsh et al., 2006). Finally, teachers also use several kinds of non-academic student data. These include attendance and student mobility data, demographic data, and data relating to students' statuses, such as whether a student has an individualized education plan (IEP) or is categorized as an English language learner in need of language supports (Gallagher et al., 2008; Marsh et al., 2006; Wayman et al., 2010).

Types of teacher data use. Research also suggests that teachers' use of data varies quite a bit, though none appears to be specific to a particular subject area. Beaver and Weinbaum (2015) propose three categories of teacher data use: Using data to inform future analysis and collection efforts; school-level responses to data analysis; and student-level interventions. They form the three categories based on an extensive review of the literature, though it should be noted that their application of the categories was limited to teachers' use of state assessment data. Teacher data use that falls into the first category includes: Setting school-wide goals, planning interim assessments, and creating school improvement plans. Data use that would fit into the second category includes: Curriculum alignment (Marsh et al., 2006), identifying teacher needs for professional development identified (Breiter & Light, 2006; Supovitz & Klein, 2003), refining course options for students (Kerr et al., 2006), and narrowing the focus of instruction to tested subjects and skills (Hamilton et al, 2005; Marsh et al., 2006). Finally, the third category of teacher data use comprises student-level interventions such as targeting specific groups of students for additional support, or identifying particular skills for which students need more practice (Booher-Jennings, 2005; Breiter & Light, 2006; Hamilton et al., 2005).

Factors influencing teacher data use. Several factors influence or shape the data teachers use, and the ways they use them. As several scholars have argued, data use is not an obvious activity;

rather, it is fraught with uncertainty because everything from the validity of the data themselves to their meaning and relevance are contested and re-negotiated through daily practice (O'Day, 2002; Spillane, 2012). For example, research suggests that across different contexts, how teachers interpret what *counts as* data can vary widely (Cho & Wayman, 2014). Indeed, the sense teachers make of data is shaped by their own beliefs and prior knowledge about the data and their students (Coburn, 2001; Coburn & Turner, 2011; Palmer & Snodgrass Rangel, 2011; Spillane et al., 2002).

What's more, teachers' sensemaking processes around data and data use are shaped by the layered contexts in which they are embedded. For example, teacher data use differs across policy (Datnow et al., 2012; Horn et al., 2015; Jennings, 2012) and district (Coburn et al., 2009; Gallagher et al., 2008; Hamilton et al., 2005) contexts. They also are shaped by the learning communities and networks in which teachers participate (Coburn, 2001; Daly, 2012; Huguet et al., 2014; Marsh et al., 2015; Nelson et al., 2008). Finally, their data use also can be influenced by school leadership, including the work of principals (Ingram et al., 2004; Wayman & Stringfield, 2006; Wayman et al., 2007, 2009) and instructional coaches (Marsh et al., 2010; Marsh et al., 2015; Swinnerton, 2007; Young, 2006).

Research has not investigated the extent to which data use varies across subject areas. Most work that has been done has focused on the highly tested areas of mathematics and reading, and has not highlighted differences that might be specific to one area or the other. Little work has set out to examine whether teachers data use varies across subject areas, though there is reason to believe that it might. The next section discusses why teacher data use might differ across subject areas, specifically in science.

Best Practices in Science Instruction and Assessment

According to the National Research Council (NRC, 2001, 2012), there are three purposes of assessment in science education: Formative assessment, summative assessment, and assessment for program evaluation. Educators (and researchers) often mistakenly rely on standardized assessments to meet the needs of all three assessment purposes. Based on its utilization of a limited number of formats (paper and pencil or computer-based) and one question type (multiple choice), this form of assessment is limited in its ability to measure performance outside of concept development, such as scientific explanations, argumentation, or the process skills associated with carrying out scientific investigations (NRC, 2012; Quellmalz, Timms, & Buckley, 2005). Best practices call for assessments to be specifically designed for the purpose of assessing content domain or the science practices of interest (NRC, 2001).

Several education policy organizations have pushed towards reforming science education by focusing not only on science content but also science practices, often through hands-on and inquiry-based instruction (Furtak, Seidel, Iverson, & Briggs, 2012; Minner et al., 2010). Inquiry-based instruction implies a different approach to assessment than traditional standardized tests. Instead of a reliance on multiple choice tests, science teachers are expected to administer a *variety* of formative and summative assessments that require students to demonstrate or perform their knowledge (Davis, Petish, & Smithey, 2006; NRC, 2000; Songer & Ruiz-Primo, 2012). These more authentic forms of assessment include labs and experiments (e.g., Hofstein & Lunetta, 2004), but also extend to classroom discussion, scientific explanations, and argumentation (Berland & McNeill, 2009; Duschl & Osborne, 2002; Duschl, Schweingruber, & Shouse, 2007; Krajcik, McNeill, & Reiser, 2008; McNeill & Krajcik, 2008; Songer & Gotwals, 2012); performance assessments such as problem- or project-based learning assignments (Kolodner et al., 2003); writing assignments (Wiggins, 1990); and classroom presentations. Utilizing a variety of assessment that is specifically designed to measure student mastery of scientific practices provides teachers with more actionable data on student mastery of both science content and scientific practices (Black & William, 1998; Buck & Trauth-Nare, 2009).

Using the research questions presented above, this study investigated the data use practices of fifth-eighth grade science teachers. Next, we explain the context of the study and present the methods used to collect and analyze the data. Then, we present the findings from the exploratory investigation. We conclude with a discussion of the findings and outline implications for research, policy, and practice in science education.

Methods

The results presented here are drawn from a broader study that focused on the development of the user interface of a PreK-12 online science curriculum (referred to below as "the study curriculum").[2] In the first year of data collection, however, the focus was on teacher data use broadly defined and the focus of the survey and our conversations with teachers was *not* on the study curriculum. The participating science teachers were asked about all of the different kinds of data they used to plan their instruction and assessments, including from other science curricula and materials. In this way, the results reported here are not about one particular curriculum, but about data use generally among science teachers. Data were collected to answer exploratory questions: Which data the science teachers use on a daily basis to plan, teach, and assess their students, and which data they *would like to use* to complete that same work.

Participants

During the summer of 2014, we recruited six suburban school districts in southeast Texas to partner on the project. Three were charter management organizations (17 schools), and three were traditional public school districts (TPS; 19 schools). The six districts constituted a convenience sample and they were chosen because they all had some experience using the study curriculum. The six districts were quite diverse (see Table 1). For example, District B, a TPS, is a suburban district that serves a majority White student population with few economically disadvantaged students and high levels of student performance in science.[3] Conversely, District A, also a suburban TPS, serves large African American and Latino student populations, with a very high percentage of their students qualifying as economically disadvantaged and lower performance on the state science assessment. The three charter school organizations (Districts D, E, and F) serve predominantly Latino and African American students, and large numbers of economically disadvantaged students. District C had the lowest passing rate in 2014 of all six of the districts, while District D had a higher passing rate than District A but a lower one than District B.

From these districts, we recruited 73 teachers, 71 of whom completed all of the research activities; 45 of them were from public schools, and 26 were from charter schools. We recruited teachers by working with the district science coordinators to email all teachers about the study; participation was voluntary and therefore the distribution of teachers was not even across the districts, and participants cannot be considered representative of other teachers in the districts.

---

[2] For an extended discussion of the study curriculum and its components that support teacher data use, see Snodgrass Rangel, Bell, & Monroy, 2016.

[3] In Texas, a student is labeled 'economically disadvantaged' if he or she qualifies for the federal free or reduced lunch program according to his or her family's income.

Table 1

Demographic data from participating districts

	District A	District B	District C	District D	District E	District F
Type of district	Traditional	Traditional	Traditional	Charter	Charter	Charter
Number of teachers interviewed	20	24	4	3	13	8
Number of campuses involved	4	11	5	2	8	8
% African American	30.5%	9.5%	5.0%	36.6%	30.0%	12.0%
% Latino	51.8%	34.1%	58.5%	60.9%	32.5%	85.0%
% Asian	3.6%	12.3%	6.2%	0.9%	13.0%	0.3%
% White	12.5%	41.0%	27.8%	0.5%	22.5%	0.8%
% Economically Disadvantaged	81.0%	29.0%	58.3%	90.7%	62.8%	86.8%
% At-risk	66.2%	37.5%	54.9%	47.2%	38.7%	37.2%
2014 5th grade Science STAAR Passing Rate	67.0%	88.0%	70.0%	63.0%	69.4%	n/a
2014 8th grade Science STAAR Passing Rate	71.0%	86.0%	73.0%	77.0%	68.9%	76.0%

Note: Data from Texas Education Agency, 2014 Performance Reports.

According to responses from our teacher survey (described in greater detail below), over half of the participating teachers had five years of experience or fewer, and over three-quarters of participating teachers had fewer than 10 years of experience (Table 2). Compared to the state average, teachers participating in this study were, on average, less experienced.

Table 2

Participants' teaching experience

Experience	Number of teachers	Teachers in study (%)	State Average (%)*
New to teaching	8	11.3	6
2 to 5 years	31	43.7	31
6 to 9 years	15	21.2	20.3
10 to 15 years	9	12.7	24.4
More than 16 years	4	5.6	18.3
Total	71	100	100

*Data from Texas Education Agency, 2010.

Teacher experience also varied across districts: Teachers in District A had significantly more experience than those in Districts E and F, while those in District C significantly more experience than teachers in District F.[4] The vast majority of teachers participating had at least some previous experience teaching science (58%), and 66.2% of all of the teachers only taught science; 25.4% taught other subjects as well, including English language arts, math, or social studies. Participating teachers were relatively evenly distributed across the four grade levels (Table 3), though there were relatively fewer teachers from sixth grade and there was one fourth grade teacher.

Table 3

Grades taught

Grades Taught	Number of teachers	%
5th	20	28.2
6th	12	16.9
7th	20	28.2
8th	22	31
Other*	3	4.2
Total	71	100

*In this category, one teacher was in 4th grade, and two taught high school in addition to middle school.

[4] Differences were significant at the $p \leq 0.05$ level.

Data Collection

We collected data from participating teachers in three ways: A teacher survey, teacher observations during planning periods or professional learning community meetings, and interviews. Here, we only report on results derived from the survey and the interviews. The survey was administered to all of the 73 teachers we recruited; 71 teachers completed the survey, for a response rate of 97%. The survey primarily drew items from existing surveys, including the Study Curriculum Evaluation Survey (Snodgrass Rangel, Bell, Monroy, & Whitaker, 2012), the Survey of Educator Data Use (Wayman, Cho, & Shaw, 2009), the School and Staffing Survey (demographic items; NCES, 2014), and the TPACK Survey (used as a measure of science content knowledge; Koehler, 2009). The survey, referenced in Appendix B and available as a supplementary file, focused primarily on general types of data that the teachers used, the frequency with which they used them, and the purposes of their data use. It included several scales, though only the data use practice scale (Cronbach's alpha = 0.95)[5] will be analyzed here. The survey did not ask the teachers what science-specific data they used, though it did provide them with opportunities to answer open-ended questions about the data and tools that would be useful for planning and collaboration. Teachers' responses to these items focused almost exclusively on the data tools they wanted to use, and not on the kinds of data they wanted to use; their responses are provided in Appendix C.

We conducted two kinds of interviews with the teachers: One-on-one interviews and focus group interviews. In total, we conducted 12 focus groups on those campuses where more than one teacher had volunteered to participate in the study; each of which included two to four teachers. The remaining teachers were interviewed one-on-one because, in all of these cases, they were the only participating teacher on their campus and the only teacher at that grade level. The interviews were semi-structured, which allowed the researchers some flexibility to address additional issues that arose during an interview that were not in the protocol (e.g., Merriam, 2009; protocol included in Appendix A). During the interviews, we were able to ask teachers about their science-specific data use. The interview protocol was developed together with the Study Curriculum developers as well as with the project's advisory board. All interviews were audio-recorded and transcribed by the research team.

Analysis of Data

We analyzed our data descriptively and analytically, both across and within district cases. Survey data were analyzed descriptively and also using independent sample *t* tests and one-way and two-way ANOVAs in order to tease out any differences across the different types of districts, the different districts, and different grade levels. The interview data were coded using a list of a priori codes that were generated from the interview protocol. All interviews were coded at least twice by different members of the research team in order to ensure reliability. Once we had coded the data, we compared results from the interviews both across grade levels and across districts and district types to identify similarities and differences in data use practices.

Limitations

As with any study, the findings presented here are limited. The first is that the focus of our data collection efforts was exploratory, with an eye towards understanding the different kinds of data that the teachers were using, as well as the kinds of data they wanted to use; the question of *how*

---

[5] Cronbach's alphas based on responses from current sample of teachers.

teachers used data and *why* came up naturally as part of our investigation, but not equally across all interviews or observations. Furthermore, our sample of districts and teachers was a convenience sample and therefore are not necessarily representative of other teachers in the districts, or of other districts. As a result, we are limited in our ability to make generalizations from our data to other teachers, schools, and districts. Furthermore, the number of teachers were able to recruit from each district was not equal, which meant that in our group comparisons, the Type I error rate may have been affected. As an example, the number of teachers per grade level was unequal as more teachers were in either fifth or eighth grades as compared to sixth or seventh grades. Another limitation inherent in any study of data use is that what counts as 'data' varies across districts and even schools. While we think that is a finding in and of itself since it reflects the way data use is framed and defined in their school and district, teachers may use data, such as from observations, that they might not consider 'data' in the formal sense and therefore may not have mentioned when prompted to elaborate on their data use. Finally, our findings are limited because they are derived from teacher self-report and from limited observations that cannot be considered representative of teachers' planning or data use practices.

Results

In this section, we present the results of our analysis of teacher data use logics. First, we examine the data teachers used in their classrooms, and then we move on to the data teachers *wanted* to use in their classrooms.

Data Science Teachers Used in their Classrooms

Here, we draw on data from the survey and interviews to describe the kinds of data the teachers used in their classrooms. In general, we found that teachers use a wide range of data to plan both their lessons and classroom assessments. We present the results both overall and also broken down by district type (charter vs. traditional public school), district, and grade level.

State assessment data. Students in Texas public schools take a science assessment three times for the purposes of both state and federal accountability, in fifth and eighth grades, and in high school for biology. Across all districts, the teachers reported relatively infrequent use of state achievement data (see Table 4; 'never' was coded as 1, and 'daily' was coded as 7). Indeed, it was the second least-frequently used kind of data and relative to the teachers' other responses, there was a fair amount of variability in responses ($M = 3.28$, SD = 1.74). There also were no significant differences in reported frequency of use of state assessment data between the charter school and traditional public school teachers. There were, however, some differences among grade levels in frequency of use of state assessment data (F=5.93[4, 62], $p < 0.05$). Specifically, fifth grade and eighth grade teachers reported using state achievement test data significantly more frequently than seventh grade teachers. These two differences make sense given that fifth and eighth grade science teachers must prepare their students for the end of year state assessment.

Table 4
Kinds of data teachers used

	State achievement data	Formal assessments (e.g., NWEA/MAP; DRI)	District-wide assessment (e.g., benchmarks)	Teacher-created assessments	Non-academic student data	Other data
Missing	2	4	2	3	3	31
Mean response	3.28	2.61	3.38	5.38	4.94	4.55
Standard deviation	1.74	1.84	1.68	1.51	1.86	2.1

There also were significant differences across districts in teachers' reported use of state assessment data. A one-way ANOVA analysis revealed that group differences existed between the following districts (F=5.41[5, 62], $p < 0.05$): Teachers in District B reported using state achievement data more frequently than teachers in District A, and teachers in District D reported using state achievement data more frequently than teachers in Districts A, C, E, and F. In District D, both the fifth and eighth grade teachers reported more frequent use of state assessment data than the teachers in the four districts named. There was, however, no significant interaction between district and grade level, suggesting that while there were differences in the frequency of use of state assessment data across districts and across grade levels, frequency of use did not depend on both district *and* grade level.

Data from the interviews confirmed what the survey uncovered, that teacher use of state achievement data was not as common as the use of other, more timely and local sources of data. Importantly, we did not ask in the interviews how frequently the teachers used each kind of data, but rather we asked them about all of the different kinds of data they used at all, and in what ways. The teachers who mentioned using state achievement data were from across all four grade levels, though most were fifth- and eighth-grade teachers, which mirrors use of these data as reported in the survey. Based on the interviews, the traditional public school teachers appeared to use state achievement data more than the teachers from the charter schools, but this difference was not supported by the survey data.

In the interviews, teachers reported using students' state assessment data to reflect, to create local assessments, and to plan long-term. For example, two fifth grade teachers at the same campus in District B described how they utilized item analyses from past state achievement tests to see what their students had struggled with in the past and to use that information to make changes to their planning with the next group of students. They also described focusing on how the questions of the state achievement tests were worded so that they could prepare future students for similar items. They told us,

> [T1]: … we use past STAAR [state achievement test data] too to kind of see what the kids have done.
> [T2]: Because it's a general trend of the same things are being missed.
> [T1]: And we look at those past STAAR questions, how are they asked, and we kind of use that to know kind of where we need to go on things that we need to make sure that we cover with them too.

Similarly, teachers reported utilizing their current students' previous state assessment data in math and reading to try to identify where students might struggle with science content and with the science assessment. For instance, an eighth grade teacher from District A told us, "we also look at their seventh grade, we talk to their seventh grade teacher to see, and we get their data to see how they did on reading and math in STAAR." This practice was particularly useful, and common, in science due to the fact that in fifth, seventh, and eighth grades, most of the teachers did not have end of year science data for their incoming students.

Teachers also described using state achievement test data as part of 'big picture' planning discussions with others in their building. These conversations often happened prior to the school year or during special planning days, when teachers, instructional coaches, and even administrators would sit and plan out whole units at a time, including how many days to dedicate to teaching each learning standard. A fifth grade teacher in District B reported that,

> …at the beginning of the school year, in some of our meetings that first week, the administrators sat with us and they had all this data, we used Lead4ward [a data reporting tool the district uses] that had compiled some of the information and we had information from our district and state, and you plug it all in to see which [Texas standards] were more challenging.

Several eighth grade teachers in District B noted that at the beginning of the school year, the past year's results were helpful for big-picture planning and looking for trends over time in student achievement, but ultimately using the previous year's eighth grade results had them comparing apples to oranges since the new students could be quite different. If the teachers wanted to look at the students' old state achievement scores, they reported that this "is hard in science because the STAAR [state achievement] data is from three years ago", and therefore not necessarily an accurate reflection of what that student has learned since fifth grade. Similarly, a sixth grade teacher in District B noted that she did not spend much time looking at students' fifth grade STAAR results because of a lack of alignment in the standards from one grade to the next.

Use of STAAR data was concentrated among the eighth grade teachers and their primary use of these data was for planning out the whole year. For example, an eighth grade teacher in District A told us that, "based on previous years, on what students struggled with the most [on the STAAR], so what they will need more time on vs. what they can pick up easily, we would dedicate a fewer number of days on that." Similarly, an eighth grade teacher from District B reported that, "I also have the released STAAR, and just an idea of how we did, what was weak, what was strong. What we need to focus on."

District assessments. Results regarding teacher use of district-wide assessments, such as interim and benchmark assessments, were consistent across both methods of data collection. In the survey, teachers responded that on average they used data from district-wide assessments between once a month and two to three times a month, and there was less variability in teachers' responses than there was for state achievement tests, suggesting that the frequency of use among teachers was more consistent (M = 3.38, SD = 1.68). When the results were disaggregated, a significant difference in the frequency of use of district-wide assessment data emerged between teachers at the charter schools and the teachers at the traditional public schools: Teachers at the charter schools on average used these data more frequently than the traditional public schools teachers (mean difference = 0.94, *t*(67) = , $p < 0.05$). There also were significant differences across districts in the frequency of district-wide data use (F=5.68[5, 62], $p < 0.05$): Teachers in District D on average reported using these data more frequently than teachers in Districts A, B, and E; and teachers in District F reported using them more frequently than teachers in Districts A, B, and E. There were no differences in the frequency of district-wide data use across grade levels.

According to our interviews with teachers, data from district-wide assessments were primarily used to to gauge students' preparation for the fifth and eighth grade STAAR assessments, and to provide interventions for those students struggling on any learning objective. It is important to note that the districts used different kinds of district-wide assessments. The traditional public school districts (A, B, and C) and one of the charter schools (District F) created their own district benchmark science assessments to test students' understanding of two to three months' worth of content. Districts A, B, and C created assessments similar to the state science assessment, while District F created its district 'common assessments' to prepare its students for the Advanced Placement exams. The traditional public schools administered the benchmarks twice a year (state law prohibits more than two administrations) to help teachers gauge whether students had learned the content during the previous months. It is likely, therefore, that data from these assessments were used more frequently than state achievement data because the district-wide assessments were administered more regularly and in every grade, so the results were more timely and relevant to all of the participating teachers.

In contrast, the three charter school districts (Districts D, E, and F) utilized the Measures of Academic Progress (MAP) interim assessments created by the Northwest Evaluation Association (NWEA). The MAP is not aligned to the Texas state standards, though it is a multiple choice assessment. Unlike the district-created assessments, the MAP assessments measure students' mastery of all of the content knowledge and skills they are expected to learn in a particular grade level. As such, the results can be used as predictive tool or an "early warning summative" tool (Wiliam & Thompson, 2007, as cited in Horn et al., 2015, p. 3).

Despite the differences in the actual tests, teachers reported using data from district-wide interim assessments in similar ways: To gauge students' preparation for end-of-year testing and to identify struggling students for interventions. For example, an eighth grade teacher in District B noted that the results from the district benchmarks "are helpful for our STAAR tutorial groups, we use those again at the end of the year." Similarly, in District A, two teams of eighth grade teachers reported using the district benchmark to track their students' progress, and also to compare their own students to those from other middle schools in the district. Several sixth and seventh grade teachers (non-tested years) told us they relied on the district benchmark data to understand their students' past performance and to do longer-term planning in lieu of using STAAR data. A fifth grade teachers in District C told us, "we need the information, I don't want to wait until the released test to see what we skipped or missed."[6] Finally, an eighth grade teacher from District F recounted, "Once we get the common assessment data, I will use that to first, see who needs a lot of help so they can do better on the next common assessment and, second, use that data towards preparation for STAAR and what groups will be separated."

We did hear about one difference from teachers in District F. They described utilizing their students' MAP scores in some cases to prepare their students for the STAAR, but in most cases to see how students were progressing towards each student's improvement goals, which the teachers created by correlating students' MAP scores to equivalent Advanced Placement (AP) scores. A

[6] As in many school districts, students in District B practiced taking a full STAAR test ahead of the real test, using one released from a previous year.

seventh grade teacher described this to us, saying, "Every kid has a growth goal for the year, which is one number higher than where they were last year, so all my 1's, my goal is to get them to be 2's this year. Is that something that realistically would happen, no but we will try to push as many as we can, and our 5's we want to stay 5's." Teachers used the correlations to predict how students might do on an AP test and to help students set their own improvement goals in relation to their readiness to take an AP test.

Local assessment data. Teachers reported using data from their own assessments more frequently than other kinds of data. In the survey, teachers responded that on average they used data from their own assessments between two to three times a week and daily; this form of data use generated the least amount of variability in teachers' responses, suggesting that the frequency of use among teachers was quite consistent (M = 5.38, SD = 1.51). There was no significant variation across grade levels, across districts, or across district types (charter vs. traditional public school districts).

Data from local assessments helped teachers in five of the six of the districts make week-to-week and even day-to-day planning decisions for their science instruction.[7] For example, teachers reported using the data from local assessments on a regular basis to create intervention or tutorial groups, particularly in the eighth grade. An eighth grade teacher in District A told us that they sat down as a team to review student scores on their twice- monthly campus assessments, nothing that these data "play a big role. We look at what our kids are struggling with the most, and like I said, we try to put those one or two questions back on the new [campus assessment] that we have, and we'll try to hit re-teaching points". Aseventh grade teacher from District F told us that at her school, "quizzes are standards-based, so I can tell which objectives we did well on or didn't do well on. That leads me to whether or not we do a full group re-take or do we do individual re-takes. Who do I pull for tutorials? Who do I pull for one-on-ones or small groups?"

Teacher use of "other data". Based on survey responses and interviews with teachers, we learned of other kinds of data and data use. Among these other kinds of data were non-academic data, which the survey defined as including student demographic data, special programmatic data (e.g., student enrolled in special education programs or is categorized as an ELL), and whether students have had disciplinary issues. On average, teachers utilized 'other' forms of data frequently, just under two to three times a week, though there was substantial variability around the mean (M = 4.94, SD = 2.1). There were no significant differences in the frequency of use of 'other' data between charter and traditional public school teachers, among grade levels, or among districts. Teachers also were encouraged to describe the other forms of data they used to understand their students and improve their instruction. Not all teachers responded, but for those who did, their responses are summarized below in Table 5. As can be seen, the teachers use a wide variety of student data in their work, some of which were science-specific (e.g., observations during in-class labs).

In the following sections, we draw exclusively on our interviews with teachers to briefly describe how teachers used 'other' data from formative assessments, classroom observations, science labs, in-class science labs, student writing, classroom discussion and questioning, performance tasks, and argumentation exercises. As is depicted in Figure 1, which summarizes the number of teachers who mentioned each data source in an interview, some of these forms of data

---

[7] Teachers in District E generally did not create their own quizzes but rather used the quizzes that came as part of the study curriculum.

were identified more frequently than others. For example, data from observations and exit tickets (formative assessment) were mentioned by more teachers than were classroom questioning/discussion, which were mentioned more frequently than student writing. Of these forms of 'other' data, some were specific to science, while others are common across subject areas.

Exit ticket data. Many teachers reported using data from a specific form of formative assessment: 'Exit tickets'. As teachers explained to us, unlike many of the other forms of assessment from which they drew data, exit tickets were not required, but rather were strongly encouraged by teachers' instructional coaches as a way to gather daily formative data on their students' learning. Exit tickets are short formative assessments that usually are administered at the end of class and ask about the concepts covered in that day's class; exit tickets would not be considered science-specific and are commonly used across subject areas. In general, teachers reported using the exit ticket data immediately, such as to make changes for the following day or week. Several teachers from across the districts told us that they use the exit ticket data to follow up in small groups to work on concepts they did not yet understand. Other teachers described using the data exit ticket data to create the next day's 'warm-up', which usually consisted of additional practice on a particular concept.

We heard about the use of exit tickets use in five of the six districts; only teachers in District E appeared not to use them. One eighth grade teacher in District F told us, "Here we use exit tickets, we are supposed to. It's a short time, usually 5 minutes of class I used exit tickets with about 3 quick questions about what we did in the class." An eighth grade teacher in District A used data from her exit tickets to make immediate adjustments to her lesson plans: "I like the exit ticket type things. Formative because then I can—Oh yes! Lesson plans have to be changed tomorrow because they didn't get this!"

Table 5

'Other' kinds of data teachers used (teacher survey)

"Other" data	Frequency
Observations of students during lab work	7
Student portfolios	1
Tracking sheets (for student learning objectives)	1
Failure reports	1
Student IEPs	2
Student writing	1
Class discussions/Questioning	2
Formative data (e.g., exit tickets)	5
In-class practice	2
Anecdotal records	2
Academic data from other classes	1
Online games	1

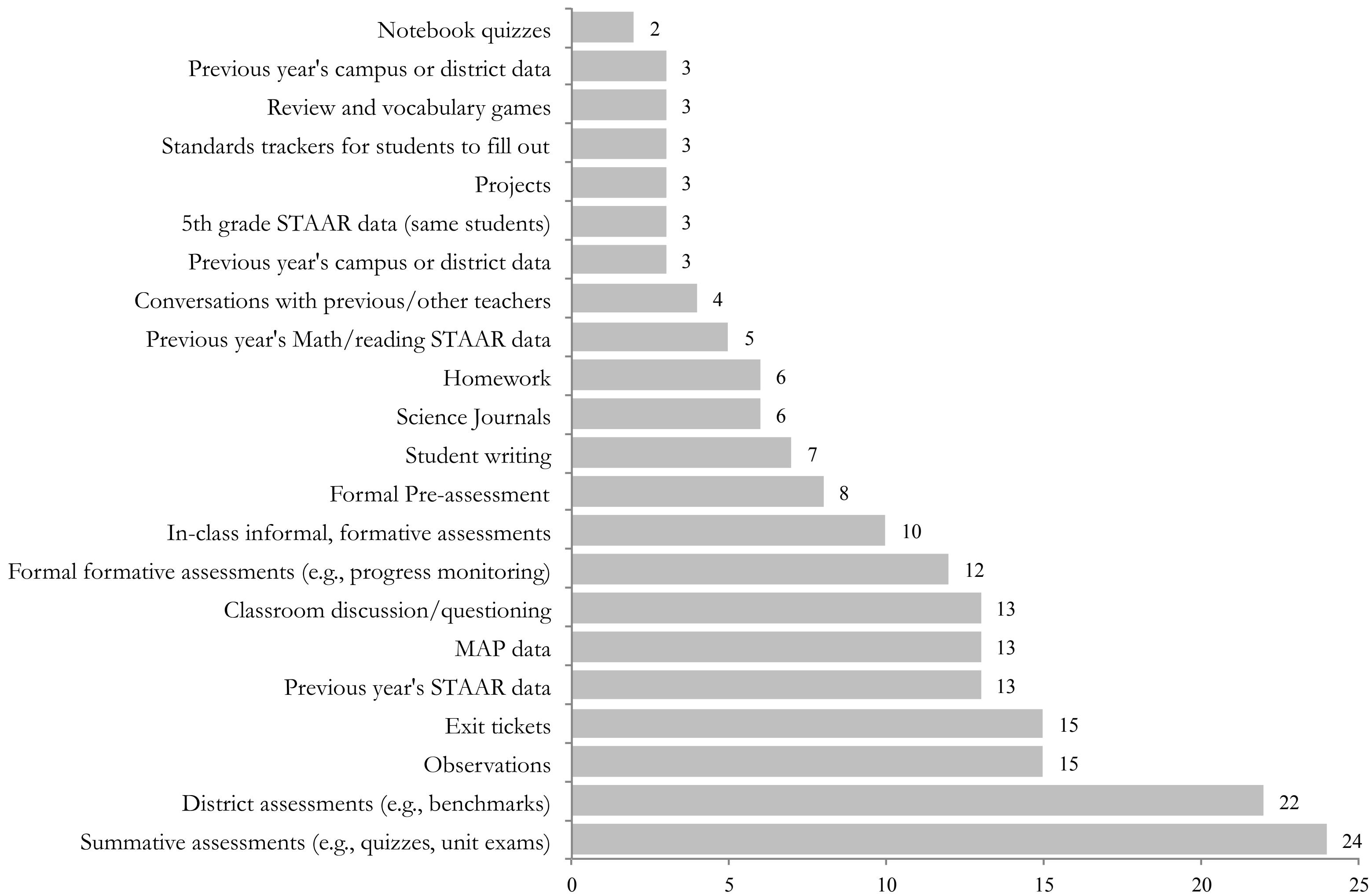

Figure 1. The number of teachers citing use of each data source (teacher interviews)

Observation data. In our interviews, several teachers discussed using their in-class observations of student work as a source of data. As with exit tickets, data from in-class observations are commonly used across subject or content areas. Many teachers from District B and District C in particular discussed using informal classroom observations to learn about what their students understood. This school district was, on average, substantially wealthier and higher performing than four of the five other districts. For example, teachers in District B described using their observations formatively to ensure that their students were understanding the material or the application of new material, and to quickly decide whether to pause a lesson to follow up with students or a certain group of students. A sixth and eighth grade teacher also in District B described observing her students and, if she saw them struggling with something, using that as a jumping off point for talking with that student's other teachers. Teacher observations appeared to go hand-in-hand with hands-on student work, such as labs and investigations. No teachers in either of the charter school districts reported use of observation data in our conversations.

The teachers also explained that they used their informal observations as a way to triangulate with other formative and summative data they already used as well as a way to see if they students really 'got it'. For instance, two fifth grade teachers recounted that, "I think we use observation a lot. Because when we do our Explore [hands-on activity], we walk around and…we get a really good sense of if they know what they're doing or if they don't." They described what they look for when they walk around and observe their students, saying, "It could just be like a deer caught in the headlights", a look that betrays a student's lack of understanding. Similarly, in a middle school, the teachers agreed that they used "a lot of it is observations" in addition to other kinds of data. Yet another team of fifth grade teachers told us they used "basic classroom observations" of students as they worked to help them decide, "what else do we need to work on, where do we need to put in more information and do things like do an extra lab."

Triangulation allowed the teachers to develop a more complete understanding of student comprehension since they were relying not on one or one type of measure, but rather two or more types of measures. Two of the fifth grade teachers in District B explained that in many cases different sources of data provided them with different, often contradictory, information about student learning. For example, one of them told us, "I think my observation, especially in science, a lot of kids who seem like they're on the ball, when you put it in a chart like this, they don't get it. Some of the kids whom I think, wow they have it, when I'm watching them, and I see it like this, and I'm shocked at some of them on the lower end of the scale [on quizzes]." As a result of the dissonance that often came of teachers' data triangulation, several of the teachers described how they had to dig deep to uncover where students' misunderstandings were—whether a student was struggling with the concept itself, or perhaps was struggling with the way a question asked about the concept.

Data from science labs. There was relatively little mention of teacher use of data derived from in-class labs or experiments, which do generate science-specific data. Indeed, in District E, the teachers did not make almost any mention of science labs, either as part of their instruction or as a source of useful data, and in District A, teachers explained that only their pre-AP students completed in-class labs. In the remaining four districts teachers reported regularly *implementing* science labs, though there was little to no discussion of using the students' work from the science labs as a source of *data*. In Districts B and F, teachers described using data from lab assessments or homework assignments that the teachers administered after the lab was completed. An eighth grade teacher in District B reported that he carefully watched students as they worked and listened to their conversations to gauge the extent to which they understood the concepts they were working with and their application. In contrast, a teacher in District F, where all science teachers were required to complete at least one lab per month, told us that district's lab assessments were not always aligned to the rest of the curriculum and therefore were not all useful as sources of data.

Data from student writing. A handful of teachers from across all four districts discussed using data from writing assignments to deepen their own understanding of what students were learning. Data derived from student writing assignments is common across other subjects and is not science-specific. The teachers who described using student writing as a source of data explained that it provided a window into students' understanding of a concept because it required them to articulate their thinking. It was not a common practice, though, in large part because of a lack of time or, as one eighth grade teacher from District F recounted, "I don't like to read a lot of writing."

An eighth grade teacher from District A explained the value that writing assignments provide to both the students and the teachers, noting that,

> More understanding of what they do and don't know from reading an essay than from a multiple choice question. Because when they have to explain it, you can tell if they really know it or not… And not just the answer, because I see that they can write the answer, but I can see that they don't understand it. If you ask them to write a sentence, you say, oh no, they know the word, but they don't really know what it is.

Similarly, an eighth grade teacher from District C explained that having his students write because writing required them to use their knowledge and therefore was a good indication of whether they had really understood the content. Finally, a teacher from District D explained that she had her students write out detailed explanations of their quiz corrections, asking them to explain "why the correct answer was correct," adding, "that helps me to know that they really got it, and put it in my notes."

Data from classroom discussions and questioning. Another informal but not science-specific way that some of the teachers described using data was through classroom discussions and questioning. Three teachers from the charter school districts D (one) and F (two) reported using both to determine if students understood the content. For instance, one middle school teacher at District F recounted how she used questioning as a way to probe for deeper understanding, describing how, "…a lot of my checks for understanding are just asking, what is your evidence for that, what is your reasoning for that?" Similarly, a middle school teacher in District D told us how she made certain to follow up with students during class as they worked on problems and she looked over their shoulders. In these cases, discussions and observations went hand-in-hand in that teachers, as they observed students work, might pose questions to students as a way to address potential student misunderstanding. In this way, interestingly, discussions often played a dual role of instructional strategy as well as formative data collection and use.

Data from performance tasks. Performance tasks require students to "actually perform, demonstrate, construct, develop a product or a solution under defined conditions and standards" (Khattri & Sweet, 1996, p. 3). While they are not unique to science, they are one way that students can demonstrate that they have mastered not just content knowledge but also scientific practices. Only two of the teachers we spoke to, one in District A and one in District B, reported using performance tasks in their classes. The eighth grade teacher in District A recalled that she had her pre-AP students complete performance tasks relatively regularly, telling us, "when we get to physics, pre-AP, they're going to do an independent project, which is to build a rollercoaster. They love it. It takes about a good day and a half. I think for regular for potential kinetic energy, we used some string and balloon and they did that activity." A seventh grade teacher in District B described one performance task that her students completed to help them understand and differentiate animal and plant cells saying, "I think we hit that, well, I know 7th grade hit that with the performance task. We did the Martian cell and they decided what kind of cell it is and explain why." The teachers did not elaborate on how they used data from these activities. Interestingly, teachers from District E, which was implementing a new project-based learning (PBL) initiative district-wide, did not discuss working on or using data from in-class projects.

Data from argumentation. Argumentation is the process by which students construct and evaluate scientific explanations for the phenomena they observe using evidence and scientific reasoning (McNeill & Krajcik, 2008), and could be considered science-specific. We only heard about the use of argumentation as an instructional strategy and as a source of data from teachers in District D. The teachers explained that, similar to student writing, argumentation was a good way to gauge the depth of students' understanding of a concept. For instance, one teacher at District D probed for deeper understanding, describing how, "…a lot of my checks for understanding are just asking, what is your evidence for that, what is your reasoning for that?" Another teacher said she was using the claim-evidence-reasoning (CER) model (Krajcik, McNeill, & Reiser, 2008; McNeill & Krajcik, 2008), which is a model for teaching students argumentation, in her classroom as a way to encourage argumentation and to ask her students to complete more rigorous work. The district was just beginning to encourage the use of argumentation and CER and this teacher told us it was hard for the students, noting, "we are now working on making claim-evidence-response/argument, but they are still struggling with that."

Data Science Teachers Want to Use in their Classrooms

While most teachers reported having the data they needed, in two districts we heard from a limited number of teachers about other kinds of data they wanted to use, or wanted to use more frequently.[8] Here we report on what those teachers told us.

Performance assessments. The teachers described performance assessments as way for students to demonstrate and even extend what they had learned in class. An eighth grade teacher from District A told us that, "In the perfect no-STAAR world, we would do instruction where [students] are actively engaged in finding the information and then have them do projects, engineering-based projects, to demonstrate their knowledge and mastery of the concept." Similarly, another group of seventh grade teachers in District A described how they wanted to give their students more projects as assessments in order to have their students apply their knowledge in a new context. Teachers also pointed out that performance assessments were more useful and engaging than multiple choice assessments because the task was more likely to resemble the real-world experiences. A teacher from District F elaborated on this feature, describing an ideal assessment as "…projects where they're building something, creating something. Something that is more real world-applicable. Because to me, when you get a job, you're not taking a test all day, you're working". Several District A teachers came back to this point, noting that students should be 'trying out' the work of scientists and engineers, applying and extending their knowledge, and not just relaying it on paper.

[8] The findings reported on in this section come exclusively from the interviews we conducted with teachers because teachers' responses to the open-ended survey questions asking about the data and tools teachers wanted to use focused almost exclusively on data *tools* and not on *kinds* of data teachers wanted to use or on the kinds of data they wanted to access through the study curriculum's dashboard.

Science labs. In addition to the performance assessments, teachers also told us they wanted to assess their students more regularly through lab assignments. The teachers saw labs as a way for students to demonstrate their knowledge as well as to 'work like scientists'. While many of the teachers we visited and spoke to described doing some kind of lab, at least a few times a semester, many teachers still felt they did not do enough labs and that they did not use the data from the labs. An eighth grade teacher in District B described the challenge of fitting in more labs because of the number of instructional days lost to local and state testing:

> And so if we could take all those days and combine them up, we'd probably get about 30 days back out of the year, 30 instructional days, not just calendar days. Thirty instructional days that we could get back, and we'd be doing some fun stuff. Using more of our chemicals that we have in the back for chemistry labs, or doing different physics projects.…

Several teachers also noted that if they had more science labs, they would have more time to observe and listen to their students, as well as ask them probing questions.

Writing assignments. Several teachers also told us that if they had more time, they would give their students more science writing assignments. The most common explanation for why teachers wanted their students to complete more writing assignments was that writing was a better source of data than most of what they were using. For example, an eighth grade teacher in District A told us that he found his students' writing revealed much about what they knew and did not know, noting that, "I'm reading carefully, if the students write really well, that means the students learned, because students are using their knowledge..." Similarly, a fifth grade teacher in District B explained why she wanted her students to write more, noting that, "Even just reading their journals and their writing you kind of can tell what they're understanding or if they're making connections with what you've been teaching or if they're not making connections."

A handful of teachers described writing in science as a way for students to both reflect and develop their own voices, and lamented that they did not have enough time to have their students write more. A 7th grade teacher from District A described this ideal, saying,

> I think if I could, I'd try to give more opportunities for story-telling, writing, because, as I'm continuing what I was saying before, they have stories to tell, but without the experience, they don't have the background knowledge to say, OK, I'm going to write about what I've done or what I know.

Another eighth grade teacher from District A noted that students in more advanced science classes were required to complete writing assignments in preparation for Advanced Placement courses in high school, but that in the 'regular' level classes, students rarely had the opportunity to write about what they were doing in class. She said, "Whereas, we've done away with essays, with the exception of Pre-AP. Where being able to express their thoughts, which is a lot more important than a multiple guess test, I guess."

Discussion

In this paper, we have described the kinds of data science teachers in grades 5 through 8 use, the ways in which they use them, and the kinds of data they would like to use more frequently.

Below, we summarize our findings and then discuss implications of the results for research and practice.

Summary of Findings

Based on the data we collected through teacher surveys and interviews, we found that this sample of science teachers used a wide range of data to plan and learn about their students. The most common data they reported using were from local assessments, including campus- and teacher-made quizzes, tests, and homework assignments. This finding was consistent across all of the schools and the grade levels. Teachers described using these local assessment data to make week-to-week and even day-to-day planning decisions for their science instruction. Teachers also told us they used data from district-wide interim or benchmark assessments. Use of these data was more frequent than data from the state assessments, though teachers at the charter schools on average used these data more frequently than the traditional public schools teachers. Teachers primarily used these data to gauge students' preparation for the fifth and eighth grade state science assessments and to provide interventions for those students struggling on any learning objective.

Teachers did use other kinds of data, and wanted to use other kinds of data, in addition to the three described in the paragraph above. According to the survey, use of other kinds of data was quite frequent, but in the survey these data were defined quite broadly. In the interviews, we did hear about other kinds of data use, including from what could be considered science-specific assessments, but it was less frequent. Most teachers we spoke to also seemed satisfied with the data they already were using, and only teachers in Districts A and B discussed other kinds of data they would like to use, such as data from science writing and labs.

Implications and Future Research

Our findings contribute to research on data use by providing additional support for how we currently understand teachers' data use, by describing the ways in which data use might be specific to a particular subject area, and by highlighting where additional research is needed. The teachers who participated in this study used data frequently in their work and they use several different kinds of data. Moreover, their use was sophisticated in that they used different kinds of data to ask and answer different questions. For example, they did not use state achievement data for daily planning, but rather for long-term planning and for reflection on their own teaching. Similarly, they used short, formative assessments for daily planning and student grouping, and not for longer-term planning. This finding supports research into how teachers alter their task or purpose according to the data they are using (e.g., Beaver & Weinbaum, 2015; Marsh et al., 2006). In other words, the teachers we surveyed and interviewed appear to use those data that are appropriate to the particular task. Future research should probe the extent to which this finding is generalizable across other teachers in other contexts, as well as whether what teachers report is represents their actual practice.

The finding that teachers use multiple sources of data also is interesting in light of the data systems to which they have access. Some research around technology suggests that it is deterministic, that is, that its design determines its use (Leonardi, 2012; Orlikowski & Iacono, 2001). The varied data use we observed, however, suggests that teachers' data use has a complex relationship with the technology meant to support it: On one hand, teachers did not rely on data from a single data management system, or even on data from *any* system, but on the other hand, they did not seem to make much use of data that could not be easily entered into a data system, such as an online grade book. For example, teachers did not talk to us about using rubrics from a performance-based task as a form of data for planning. In other words, we did not hear about data use that was determined by technology, but rather facilitated by it. Future research should explore more carefully the evolving relationship between teachers' data use and the technology that is meant to support it.

In contrast to accounts of data use where teachers reported frustration with the push to use what many considered to be untimely or limited state achievement data (e.g., Beaver & Weinbaum, 2015; Ingram et al., 2004), we did not hear consistent frustration from teachers. Rather, all of the teachers demonstrated an understanding that state achievement data were one part of their broader data use practices, which also included district and local assessments. This is not to say that some teachers were not frustrated with what in some districts appeared to be a narrow focus on using data from the state achievement tests: Teachers in District A in particular expressed disapproval for how the achievement tests and preparation or the tests had crowded other kinds of assessments that might yield richer data about their students' learning in science. Additional research will be needed to understand the extent of this frustration among teachers, as well as the ways in which this frustration is distributed across teachers, schools, and districts. Given that District A had the lowest pass rates during the two years prior to this study, it is plausible to hypothesize that they are under more or different pressure than the teachers in the other districts where students had higher performance; the design of this study did not allow us to link student achievement to data use practices directly, but the data suggest that this hypothesis is worth testing.

We found relatively limited use of science-specific assessment data. As we outlined in our literature review, we expected that we would observe data use that extended beyond what the literature had identified as common 'generic' data use practices across different teachers. Among the assessments from which we hypothesized teachers would use data were hands-on activities and labs, performance tasks, and argumentation activities. In the interviews, a small number of teachers described completing these activities, but either did not mention using data from them or explicitly noted that they did not use data from these activities. It is not clear from this study why the teachers either did not implement these science-specific assessments that are considered best practices within the science education community or did not use data from the science-specific assessments they did administer in their classrooms. Based on previous research into the implementation of inquiry-based science (e.g., Crawford, 2007; Davis et al., 2006; Johnson, 2006), it is possible to speculate that teachers continue to struggle with inquiry in their classrooms and, for this reason, do not have data from inquiry-based assessments, or perhaps are not sure about how to use data from inquiry-based assessments. Other research suggests that novice teachers in particular have relatively unsophisticated knowledge of science content and processes as well as the nature of science, which perhaps would make them uncomfortable with or unprepared for hands-on, inquiry-based instruction and or for anything but multiple choice assessments (Davis et al., 2006). There also is research that would support speculation that a narrow focus on using data from the state assessment and other assessments designed to prepare for the state assessment simply may crowd out more authentic assessment, whether in science (Marx & Harris, 2006) or in other areas (e.g., Palmer & Snodgrass Rangel, 2011). Future research will need to examine more closely the full chain of teacher data use, including assessment, data analysis, and instructional planning and decision making to understand how teachers utilize science-specific assessments and activities and, for those who do not, why.

While the nature of our sample makes it difficult to generalize to other teachers, our findings do point to potential areas of improvement for practitioners. The results suggest that teachers take seriously the mandate to use data regularly to improve their instruction and long-term planning, but they also indicate that their data use could be broadened to incorporate data from science-specific assessments. On one hand, this is a problem of teachers not administering these assessments on a regular basis, but on the other hand our findings suggest that it may be a problem of helping teachers *incorporate* data from these assessments when they are administered.

Finally, future research on data use among science teachers should address the central limitation of this study: The research design. The teachers who participated in this study did not comprise a random sample and therefore cannot be considered representative of the population of teachers in any of the six districts. For this reason, it is not possible to generalize the findings beyond the teachers we interviewed. A future study would benefit from a comparative design that investigates the extent to which practices are specific to districts. Further, the districts should be selected with an eye towards representing diverse contexts: Rural, urban, suburban, traditional public, public charter, and private schools. From there, a stratified random sampling approach would be a more fruitful strategy for selecting specific schools and teachers. Finally, a more comprehensive data collection strategy combining survey responses, one-on-one interviews, planning observations, and classroom observations would yield richer data. Such a design would permit an investigation of the extent to which the key findings presented here are unique to the contexts and teachers we studied, or are common to science teachers more broadly.

Acknowledgements

This material is based upon work supported, in part, by the National Science Foundation under Grant No. 1417705.

The second author works for the publishing company that owns and produces the science curriculum that was the subject of the broader study reported on in part here.

None of the authors of this article contributed to the creation of the curriculum that was the focus of the NSF-funded study.

References

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching*, *47*(5), 564-582.

Auger, R. W., & Blackhurst, A. E. (2005). The development of elementary-aged children's career aspirations and expectations. *American School Counselor Association Journal, 8*(4), 322-329.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. 2001. Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development, 72*(91), 187-206. http://dx.doi.org/10.1111/1467-8624.00273

Beaver, J. K. & Weinbaum, E. H. (2015). State test data and school improvement efforts. *Educational Policy, 29(*3), 478-503. http://dx.doi.org/10.1177/0895904813510774

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94*(5), 765-793. http://dx.doi.org/10.1002/sce.20402

Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74. http://dx.doi.org/10.1080/0969595980050102

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System. *American Educational Research Journal, 42*(2), 231-268. http://dx.doi.org/10.3102/00028312042002231

Breiter, A., & Light, D. (2006). Data for school improvement: Factors for designing effective information systems to support decision-making in schools. *Journal of Educational Technology & Society*, *9*(3), 206-217.

Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching, 47*(4), 402-421. http://dx.doi.org/10.1002/tea.20344

Carraway, J. H., & Young, T. (2015). Implementation of a Districtwide Policy to Improve Principals' Instructional Leadership Principals' Sensemaking of the Skillful Observation and Coaching Laboratory. *Educational Policy*, *29*(1), 230-256. http://dx.doi.org/10.1177/0895904814564216

Cho, V., & Wayman, J. C. (2014). Districts' efforts for data use and computer data systems: The role of sensemaking in system use and implementation. *Teachers College Record*, *116*(2), 1-45.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145-170. http://dx.doi.org/10.3102/01623737023002145

Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational policy*, *19*(3), 476-509. http://dx.doi.org/10.1177/0895904805276143

Coburn, C. E., Touré, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *The Teachers College Record*, *111*(4), 1115-1161.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, *9*(4), 173-206. http://dx.doi.org/10.1080/15366367.2011.626729

Copland, M. A. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis*, *25*(4), 375-395. http://dx.doi.org/10.3102/01623737025004375

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, *44*(4), 613-642. http://dx.doi.org/10.1002/tea.20157

Daly, A. J. (2012). Data, dyads, and dynamics: Exploring data use and social networks in educational improvement. *Teachers College Record*, *114*(11), 1-38.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2012). High school teachers' use of data to inform instruction. *Journal of Education for Students Placed at Risk (JESPAR), 17*(4), 247-265. http://dx.doi.org/10.1080/10824669.2012.718944

Datnow, A., Park, V., & Wohlstetter, P. (2007). Achieving with data: How high-performing school systems use data to improve instruction for elementary students. Los Angeles: *USC Center on Educational Governance*. Retrieved from http://www.usc.edu/dept/education/cegov/achieving_data.pdf

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*(4), 607-651. http://dx.doi.org/10.3102/00346543076004607

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*(1), 39-72. http://dx.doi.org/10.1080/03057260208560187

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, DC: National Academies Press.

Fuhrman, S., & Elmore, R. F. (Eds.). (2004). *Redesigning accountability systems for education*. New York: Teachers College Press.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300-329. http://dx.doi.org/10.3102/0034654312457206

Gallagher, L., Means, B., & Padilla, C. (2008). *Teachers' Use of Student Data Systems to Improve Instruction: 2005 to 2007*. Washington, DC: US Department of Education.

Hamilton, L. S., Berends, M., & Stecher, B. M. (2005). *Teachers' responses to standards-based accountability*. Santa Monica, CA: RAND Corporation.

Hamilton, L. S., & Stecher, B. (2006). Measuring instructional responses to standards-based accountability. Santa Monica, CA: RAND Corporation.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: Rand Corporation.

Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies. *The Economic Journal, 113*(485), F64-F98. http://dx.doi.org/10.1111/1468-0297.00099

Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management, 24(*2), 297-327. http://dx.doi.org/10.1002/pam.20091

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28-54. http://dx.doi.org/10.1002/sce.10106

Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making Sense of Student Performance Data: Data Use Logics and Mathematics Teachers' Learning Opportunities. *American Educational Research Journal*. doi: 0002831215573773.

Huguet, A., Marsh, J. A., & Farrell, C. (2014). Building teachers' data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives*, *22*(52), 1-31. http://dx.doi.org/10.14507/epaa.v22n52.2014

Ingram, D., Seashore Louis, K., & Schroeder, R. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, *106*(6), 1258-1287. http://dx.doi.org/10.1111/j.1467-9620.2004.00379.x

Jennings, J. L. (2012). The effects of accountability system design on teachers' use of test score data. *Teachers College Record*, *114*(11), 1-23.

Jimerson, J. B. & Wayman, J. C. (2015). Professional learning for using data: Examining teacher needs and supports. *Teachers College Record, 117*(4), 1-36.

Johnson, C. C. (2006). Effective professional development and change in practice: Barriers science teachers encounter and implications for reform. *School Science and Mathematics*, *106*(3), 150-161. http://dx.doi.org/10.1111/j.1949-8594.2006.tb18172.x

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons fromthree urban districts. *American Journal of Education*, *112*(4), 496-520. http://dx.doi.org/10.1086/505057

Khattri, N., & Sweet, D. (1996). Assessment reform: Promises and challenges. In M. B. Kane & R. Mitchell (Eds.), *Implementing performance assessment* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., . . . Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design (tm) into practice. *The Journal of the Learning Sciences, 12*(4), 495-547. http://dx.doi.org/10.1207/S15327809JLS1204_2

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education, 92*(1), 1-32. http://dx.doi.org/10.1002/sce.20240

Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk, 10*(3), 333-349. http://dx.doi.org/10.1207/s15327671espr1003_7
Leonardi, P. M. (2012). Materiality, sociomateriality, and socio-technical systems: What do these terms mean? How are they related? Do we need them? In P. M. Leonardi, B. A. Nardi, & J. Kallinikos (Eds.), *Materiality and Organizing: Social Interaction in a Technological World* (pp. 25–48). Oxford: Oxford University Press.
Long, L., Rivas, L., Light, D., & Mandinach, E. B. (2008). *Data-driven school improvement: Linking data and learning.* New York: Teachers College Press.
Marsh, J. A., Bertrand, M., & Fuguet, A. (2015). Using data to alter instructional practice: The mediating role of coaches and professional learning communities. *Teachers College Record, 117*, 1-41.
Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How Instructional Coaches Support Data-Driven Decision Making Policy Implementation and Effects in Florida Middle Schools. *Educational Policy*, *24*(6), 872-907. http://dx.doi.org/10.1177/0895904809341467
Marsh, J. A., Pane, J. F., & Hamilton, S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research*. Santa Monica, CA: RAND Corporation.
Marx, R. W., & Harris, C. J. (2006). No child left behind and science education: Opportunities, challenges, and risks. *Elementary School Journal, 106*, 455–466. http://dx.doi.org/10.1086/505441
McLaughlin, M., & Mitra, D. (2003). *The cycle of inquiry as the engine of school reform: Lessons from the Bay Area School Reform Collaborative*. San Francisco, CA: Bay Area School Reform Collaborative.
McNeill, K. L., & Krajcik, J. (2008). Inquiry and scientific explanations: Helping students use evidence and reasoning. In J. Luft, R. Bell, & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 121 – 134). Arlington, VA: National Science Teachers Association Press.
Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education.
Merriam, S. B. (2009). *Qualitative Research. A guide to Design and Implementation: Revised and Expanded from Qualitative Research and Case Study Applications in Education*. San Francisco, CA: Jossey-Bass.
Milner, A. R., Sondergeld, T. A., Demir, A., Johnson, C. C., & Czerniak, C. M. (2012). Elementary teachers' beliefs about teaching science and classroom practice: An examination of pre/post NCLB testing in science. *Journal of Science Teacher Education, 23(*2), 111-132. http://dx.doi.org/10.1007/s10972-011-9230-7
Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474-496. http://dx.doi.org/10.1002/tea.20347
National Center for Education Statistics (NCES). (2011). The Nation's Report Card: Science 2011. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2011/2012465.pdf.
National Center for Education Statistics. (2014). *School and staffing survey*. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/surveys/sass/
National Research Council. (2000). *Inquiry and the national science education standards*. Washington, DC: National Academies Press.
National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

National Research Council (2011). *Successful K-12 STEM Education: Identifying effective approaches in science, technology, engineering, and mathematics.* Washington, DC: National Academy Press.
National Research Council (2012). *A framework for K-12 science education*. Washington, DC: National Academy Press.
Nelson, T., Slavit, D., Perkins, M., & Hathorn, T. (2008). A culture of collaborative inquiry: Learning to develop and support professional learning communities. *The Teachers College Record*, *110*(6), 1269-1303.
O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review, 72*(3), 293-329. http://dx.doi.org/10.17763/haer.72.3.021q742t8182h238
O'Day, J. A., & Smith, M.S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system,* (pp. 250-312). San Francisco: Jossey-Bass.
Opfer, V. D., Henry, G. T., & Mashburn, A. J. (2008). The district effect: Systemic responses to high stakes accountability policies in six southern states. *American Journal of Education, 114*(2), 299-332. http://dx.doi.org/10.1086/521242
Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the "IT" in IT research - A call to theorizing the IT artifact. *Information Systems Research*, *12*(2), 121–135. http://dx.doi.org/10.1287/isre.12.2.121.9700
Palmer, D. & Snodgrass Rangel, V. W. (2011, July). High stakes accountability and policy implementation: Teacher decision-making in bilingual classrooms in Texas. *Educational Policy, 25*(54), 614-647. http://dx.doi.org/10.1177/0895904810374848
Quellmalz, E. S., Timms, M., & Buckley, B. (2009). *Using science simulations to support powerful formative assessments of complex science learning.* Paper presented at the annual meeting of the American Educational Research Association. San Diego, CA.
Rangel, V. S., Bell, E. R., & Monroy, C. (2016). *Working with middle school science teachers to design and implement an interactive data dashboard.* Presentation at the annual meeting of the American Educational Research Association.
Rangel, V. S., Bell, E. R., Monroy, C., & Whitaker, J. R. (2012). *STEMscopes Evaluation Survey.* Houston, TX: Rice University.
Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education*, *95*(3), 458-476. http://dx.doi.org/10.1002/sce.20431
Schmidt, D. A., Baran, E., Thompson, A. D, Koehler, M. J., Mishra, P., & Shin, T. (2009). *Survey of Preservice Teachers' Knowledge of Teaching and Technology.* http://mkoehler.educ.msu.edu/unprotected_readings/TPACK_Survey/tpack_survey_v1point1.pdf
Smith, M. S. & O'Day, J. (1990). Systemic school reform. *Journal of Education Policy, 5*(5), 233-267. http://dx.doi.org/10.1080/02680939008549074
Songer, N. B., & Gotwals, A. W. (2012). Guiding explanation construction by children at the entry points of learning progressions. *Journal of Research in Science Teaching, 49(*2), 141-165. http://dx.doi.org/10.1002/tea.20454
Songer, N. B., & Ruiz-Primo, M. A. (2012). Assessment and science education: Our essential new priority? *Journal of Research in Science Teaching, 49*(6), 683-690. http://dx.doi.org/10.1002/tea.21033
Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education*, *118*(2), 113-141. http://dx.doi.org/10.1086/663283

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387-431. http://dx.doi.org/10.3102/00346543072003387

Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change, 10*(2-3), 211-227. http://dx.doi.org/10.1007/s10833-009-9105-2

Supovitz, J. (2012). Getting at student understanding–The key to teachers' use of test data. *Teachers College Record, 114*(11), 1-29.

Supovitz, J., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools use student performance data to guide improvement*. Philadelphia: Consortium for Policy Research in Education.

Swinnerton, J. (2007). Brokers and boundary crossers in an urban school district: Understanding central-office coaches as instructional leaders. *Journal of School Leadership*, *17*(2), 195.

Tai, R., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*, 1143-1144. http://dx.doi.org/10.1126/science.1128690

Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, *44*(3), 519-558. http://dx.doi.org/10.3102/0002831207306859

Wayman, J. C. (2005). Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at* http://dx.doi.org/10.14507/epaa.v20n25.2012

Wayman, J. C., & Jimerson, J. B. (2014). Teacher needs for data-related professional learning. *Studies in Educational Evaluation*, *42*, 25-34. http://dx.doi.org/10.1016/j.stueduc.2013.11.001

Wayman, J. C., Cho, V., Jimerson, J. B. & Snodgrass Rangel, V. W. (2015). A look into the workings of data use in a mid-sized district. In Sutherland, I., Sanzo, K. S., & Scribner, J. P. (Eds.) *Leading small and mid-sized urban school districts.* United Kingdom: Emerald Books. *Risk, 10*(3), 295-308. http://dx.doi.org/10.1207/s15327671espr1003_5

Wayman, J. C., Cho, V., Jimerson, J. B., & Spikes, D. D. (2012). District-wide effects on data use in the classroom. *Education Policy Analysis Archives*, *20*(25). Wayman, J. C., Cho, V., & Johnston, M. T. (2007). *The data-informed district: A district-wide evaluation of data use in the Natrona County School District.* Austin, TX: The University of Texas at Austin.

Wayman, J. C., Cho, V., & Shaw, S. (2009). *Survey of educator data use.* Austin, TX: University of Texas at Austin.Wiggins, G. (1990). The Case for Authentic Assessment. ERIC Digest.

Wayman, J. C., Snodgrass Rangel, V. W., Jimerson, J. B., Cho, V., (2010). *Improving Data Use in District Three: Becoming a Data-Informed District.* Austin: The University of Texas.

Wayman, J. C., & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education*, *112*(4), 549-571. http://dx.doi.org/10.1086/505059

Wayman, J. C., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through the analysis of student data* (Report No. 67). Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

Wiggins, G. (1990). *The case for authentic assessment*. Washington, DC: ERIC Clearinghouse on Tests, Measurement, and Evaluation.

Young, V. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American Journal of Education 112*(4), 521–48. http://dx.doi.org/10.1086/505058

Appendix A
Teacher Focus Group Protocol

Introduction:
My name is ___ and I work at - University. We're working on a research project sponsored by the National Science Foundation that aims to understand how you use data and reports to plan for science instruction, to collaborate with your colleagues, and to improve your instruction so that we can work with the [Study] science curriculum to improve the teacher dashboard. We appreciate your willingness to meet with us. The focus group interview should take 45 minutes to an hour, though you can stop the focus group at any time. Everything you say here will remain anonymous and we'll never identify your or school in any reports. I would like to record the interview, is that OK? I can turn the recorder off at any point if you would like. (Give teachers consent forms to sign)

I. *We're going to start with some general questions about how you use data to plan, assess, and think about your instruction.*
   1. In a world with no accountability and high stakes testing, how do you think you would figure out what your students know and don't know in science, and decide what to do about it?
   2. Tell me about how you plan to teach science:
      a. Do you plan alone, collaboratively, or both?
      b. What do you have with you when you plan?
   3. What data do you consider important to help you teach science?
      a. Why do you think that's the most useful data?
      b. Where do you get that data from?
      c. What else would you like to have?
   4. What kind of information might lead you to make a change in how you're teaching science?
      a. Can you give an example of a change you've made, and why you made it?

II. *Now, we're going to switch gears and ask you some questions about other resources you use in any of the courses you teach. For all of these questions, please think about other digital curricula and tools that you use to plan, teach, and assess.*
   1. What other digital resources or tools do you use in your classes (e.g. Edmodo, Eduphoria, district-created system, etc.)?
   2. When you think about the teacher interface or dashboards in those systems, what features and tools make them easier or harder to use?

III. *Now we're going to ask a few questions that are specific to the [Study] curriculum.*
   1. What [Study] curriculum data tools do you use?
   2. How do these tools compare to what's available in other programs or digital tools that you use?
   3. If you could design an online tool to help you do what you need to do in your science classes, what would it have? What would it help you do?

Is there anything else you would like to add about how the [Study] curriculum plays a part of your science planning and teaching, or anything else we've talked about here?

Appendix B: Teacher Survey

Full-text of the teacher survey is available as a separate supplementary file.

The survey primarily drew items from existing surveys, including the Study Curriculum Evaluation Survey (Snodgrass Rangel, Bell, Monroy, & Whitaker, 2012), the Survey of Educator Data Use (Wayman, Cho, & Shaw, 2009), the School and Staffing Survey (demographic items; NCES, 2014), and the TPACK Survey (used as a measure of science content knowledge; Koehler, 2009). The survey focused primarily on general types of data that the teachers used, the frequency with which they used them, and the purposes of their data use. It included several scales, though only the data use practice scale (Cronbach's alpha = 0.95)[9] is a focus of this study. The survey did not ask the teachers what science-specific data they used, though it did provide them with opportunities to answer open-ended questions about the data and tools that would be useful for planning and collaboration.

---

[9] Cronbach's alphas based on responses from current sample of teachers.

Appendix C: Teachers' Open-ended Survey Responses

What other tools or reports or kinds of data/information would help you plan your lessons?	What other tools or reports or data/information would help you meet with other teachers or instructional coaches to talk about science instruction or learning?
Beginning of year Science Assessment (information from data)	Assessment Data
A bigger science question bank and test builder- one that is searchable by TEK.	Use eduphoria data tables from test or quizzes
District student data and application of our content.	We use district benchmark data, as well as informal assessments.
eduphoria	We used the quizzes, some of the student journal activites, vocabulary, OER's and reading selections
How to tie in the TEKS explicitly.	I don't meet with other teachers to plan science instruction.
I like the idea of a item analysis. I am going to look for where that is on stemscopes	I think more rational for the use of data and standards-based grading would be beneficial for teachers to learn to encourage its use.
I think what is available is great, I just need more training and a refresher.	I think what is available is great. I just need more training and a refresher.
I used the quizzes, vocabulary activities, intervention activities, TEKS unwrapped, OER's, readings, and student journal activities for some of the units.	I would need more training on the full use of stemscopes. We are currently just pulling activities and assessments from stemscopes.
I would love it if STEMscopes color coded the students by their mastery level of the completed activity (similar to Kickboard).	Maybe pbl
I'd like to learn more about the STEMscopes tools as I have not heard of them before.	NWEA data, Map testing, RIT scores.
I'm not sure.	pre-assessments
Lab safety posters, assessments	Pre-post assessment data
Maybe some hands on activities to engage students.	Stemscopes, videos, ppt
More levels of activities so that I could differentiate for co-taught classes.	TEKS and more labs

One central place that allows me to keep track of all of my data and also ties into the gradebook so that I'm not exporting data all the time.	The item analysis feature of Eduphoria is useful in assessing what question was missed by most students and which TEKS need to be re-taught in order to reinforce the concepts.
Simple practice questions and guided practice activities	
Sometimes it would help to have an idea of how many minutes/periods (of 45 minutes) a particular activity may take	
Stemscopes, videos, ppt	
STAAR data, district curriculum resources	

About the Authors

Virginia Snodgrass Rangel
University of Houston
Vrangel3@uh.edu
http://orcid.org/0000-0002-0376-1986
Dr. Snodgrass Rangel is an Assistant Professor in the Department of Educational Leadership and Policy Studies at the University of Houston. Her research interests include teacher leadership, data use policy and practice, policy implementation, school reform, and STEM education. Her work has appeared in the *American Educational Research Journal, Educational Policy, and the Journal of School Leadership.*

Carlos Monroy
Rice University
Carlos.monroy@rice.edu
Dr. Monroy is a Research Scientist in the Computer Science Department at Rice University. He works with the Pliny Project, which seeks to gather hundreds of billions of lines of publicly available open-source computer code, and to mine that code to create a searchable database of properties, behaviors, and vulnerabilities. His research interests include data mining, information retrieval and visualization, and learning analytics. The goal of his work is to harness data science methods for fostering innovation, create meaningful contexts, and improve knowledge discovery.

Elizabeth Bell
Accelerate Learning Inc.
ebell@acceleratelearning.com
Dr. Bell is the Research Manager at Accelerate Learning Inc. Her research interests include social and cognitive development as it applies to early childhood education as well as the implementation and impact of science curriculum in diverse classrooms.

education policy analysis archives

Volume 24 Number 86 August 8. 2016 ISSN 1068-2341