Contemporary Educational Psychology 57 (2019) 9–31

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Why are we learning this? Using mixed methods to understand teachers' relevance statements and how they shape middle school students' perceptions of science utility

Jennifer A. Schmidt[a,*], Stephen S. Kafkas[b], Kimberly S. Maier[c], Lee Shumow[b], Hayal Z. Kackar-Cam[b]

[a] Michigan State University, Erickson Hall, 620 Farm Lane Room 513D, East Lansing, MI 48824, USA
[b] Northern Illinois University, Department of Leadership, Educational Psychology & Foundations, Graham Hall, DeKalb, IL 60115, USA
[c] Michigan State University, Erickson Hall, 620 Farm Lane Room 451, East Lansing, MI 48824, USA

ARTICLE INFO

Keywords:
Instructional relevance
Utility value
Science
Mixed methods
Middle school

ABSTRACT

We investigated how four middle school science teachers perceived and communicated the relevance of science content in their seventh-grade classrooms (n = 14), and examined their students' (N = 306) perceptions regarding the utility of daily course content and the domain of science more generally. Teacher interviews and repeated classroom observations were used to build an understanding of how and for whom science teachers saw science content as relevant, and to examine the different ways teachers communicated content relevance to students during instruction. Students' perceptions of science utility were measured using repeated self-reports of the usefulness of daily content and traditional surveys assessing global science utility. Following separate analyses of the quantitative and qualitative data strands, the data sources were merged to investigate how teachers' relevance statements relate to their own stated beliefs, and to students' daily and global perceptions of science utility. Teachers varied in terms of the depth and breadth of connections they saw between science content and students' lives, and these beliefs were largely reflected in the relevance statements they made while they were teaching. Students perceived the domain of science as being moderately useful, and often did not see the utility of their daily course activities. When teachers made more frequent relevance statements, their students were more likely to see daily content as useful and had increased perceptions of global science utility. Results highlight the influence of teacher beliefs on practice, as well as the immediate and longer-term effects of relevance strategies on students' utility beliefs.

1. Introduction

"Why do I have to learn this?" "When will I ever need to know this?" Such questions are all too familiar to teachers, who are regularly tasked with facilitating learning of academic content among students who perceive that content as having little meaning, applicability, or value beyond the course in which it is introduced. Many students view schoolwork as irrelevant, or "unrelated to issues, competencies, or concerns of the real world" (Newmann, Wehlage & Lamborn, 1992, p. 26), and researchers, educational reformers, instructional designers, and students themselves have long pointed to this lack of relevance in explaining low levels of student engagement and motivation (Keller, 1987a; National Research Council, 2012; NGSS Lead States, 2013; Newmann et al., 1992).

Theory and research suggest that when learning tasks have meaning, applicability or value to students outside the instructional context, students exhibit greater interest, engagement, performance, and persistence (Eccles et al., 1983; Eccles, Barber, Updegraff & O'Brien, 1998; Keller, 1987a; Hulleman & Harackiewicz, 2009; Hulleman, Durik, Schweigert & Harackiewicz, 2008; Lau & Roeser, 2002; Newmann et al., 1992; Simpkins, Davis-Kean & Eccles, 2006). The degree to which students perceive particular academic tasks or course subjects as meaningful is undoubtedly influenced by a variety of factors that students bring with them to class like interests, life experiences, current understandings, goals, cultural norms, and stereotypes. These perceptions are also likely shaped by teachers and the extent to which they present their course material as having relevance outside the classroom. We know relatively little about how teachers

* Corresponding author.
E-mail addresses: jaschmid@msu.edu (J.A. Schmidt), kmaier@msu.edu (K.S. Maier), hkackar@niu.edu (H.Z. Kackar-Cam).

https://doi.org/10.1016/j.cedpsych.2018.08.005

Available online 22 August 2018
0361-476X/

attempt to communicate relevance during everyday instruction or whether these efforts have any real impact on how students view their schoolwork.

Effectively communicating relevance through instruction may be especially important in the domain of science because scientific knowledge and reasoning can inform one's everyday choices about diet, health, and purchasing, and can facilitate understanding of contemporary sociopolitical issues such as those related to the environment and energy (Schreiner & Sjøberg, 2004). Additionally, many jobs that provide a living wage (including jobs for which a college education is not necessary) require some degree of scientific literacy (Rothwell, 2013; Vilorio, 2014). For these reasons, the framework underlying the Next Generation Science Standards highlights the importance of making science content relevant (NGSS Lead States, 2013; National Research Council, 2012).

It may be particularly important to understand how middle school teachers frame science relevance. Students' interest in, motivation for, and valuing of academic subjects tends to decline throughout the middle school years (Gnambs & Hanfstingl, 2016; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002), and such declines are especially precipitous in mathematics and science (Greenfield, 1997; George, 2000; Gottfried, Fleming, & Gottfried, 2001; Osborne, Simon, & Collins, 2003). Identifying ways middle school teachers can effectively help students see the importance of their science activities might suggest pathways for stemming the motivational declines observed during this developmental period.

In this article we use qualitative and quantitative methods to provide descriptive accounts of how four seventh-grade science teachers thought about and communicated the relevance of their subject matter. We then explore relations among teachers' expressions of science relevance during the course of their naturally occurring instruction to students' beliefs about the value of their immediate science content and the domain of science more generally.

1.1. Relevance as an instructional strategy

In education literature, the term "relevance" is often used but seldom defined. Drawing upon theory and research in instructional design and technology (Clegg & Kolodner, 2014; Keller, 1987a; Means, Jonassen & Dwyer, 1997), instructional communication (Frymier & Shulman, 1995; Muddiman & Frymier, 2009), and educational psychology (Anderman, Andrzejewski & Allen, 2011; Assor, Kaplan & Roth, 2002; Bergin, 1999; Newmann et al., 1992; Skinner & Belmont, 1993), we use the term relevance to refer to instructional strategies that teachers use to highlight the meaning, applicability or usefulness of course content beyond the immediate instructional context. That is, we use relevance to refer to instructional strategies that are intended to connect their content to issues, competencies and concerns of the "real world." As such, relevance is a broad term that includes targeted connections teachers make to students' immediate or goals and interests (Skinner & Belmont, 1993) as well as less specific connections between course content and life outside the classroom (e.g., local environments, current events; see Clegg & Kolodner, 2014; Frymier & Shulman, 1995; Keller, 1987a; Means et al., 1997; Newmann et al., 1992).

Research on effective teaching suggests that students tend to perceive their classroom environment as more motivating, have more positive attitudes toward school, and engage more deeply when they feel their teachers use relevance in their instruction (Anderman et al., 2011; Assor, et al., 2002; Bergin, 1999). Clegg and colleagues have done instructional design work in which they create "life-relevant science-learning programs" where learners connect science to their everyday lives. The researchers argue that through exposure to these everyday connections, students develop scientific dispositions and see the value of science in relation to their goals (Clegg et al., 2014; Clegg & Kolodner, 2014). While making content relevant is generally recognized as good teaching practice, we know little about how teachers in general or science teachers in particular think about the relevance of their content. In what ways do science teachers view their content as related to life outside the classroom? Do they see connections for all students or just some students? Do they see connections to their students' future and/or to their daily experience? Teachers may differ widely in the breadth and depth of connections they see, and one goal in the current study is to understand how and for whom the science teachers in our study perceived the relevance of their content.

Research on the effects of teachers' use of relevance has typically relied on student reports about whether they perceive their teachers as promoting relevance, which are then linked to students' reports of motivation and value (Assor et al., 2002; Frymier & Shulman, 1995; Schreiner & Sjøberg, 2004). It is potentially problematic to rely on student reports of teacher behaviors because students who are more motivated and value the course content more to begin with might see more relevance in their teachers' behaviors. Instructional design research often incorporates relevance information through instructional texts rather than implementation of specific teacher behaviors (Means et al., 1997), or does not systematically document how teachers are actually implementing learned relevance strategies in the classroom (Keller, 1987b). How teachers communicate relevance to their students is likely shaped by teachers' content knowledge, by the degree to which they see their content as relevant, and by their perceptions of their students. Descriptive details about how teachers communicate relevance during instruction might help us understand whether and how the use of relevance is related to students' beliefs and behaviors. To this end, a second goal of this study is to understand the different ways teachers communicate relevance during instruction.

1.2. Perceived utility as the means through which relevance "works"

Relevance strategies are presumed to benefit student motivation and learning because they help students personally value what they are learning and see it as related to the fulfillment of their goals (Assor et al., 2002; Clegg & Kolodner, 2014; Keller, 1987a). As a result of this increased valuing, students are more willing to focus attention and effort on coursework that they see as worthwhile, which results in increased engagement and learning. Thus, perceived value is a means through which relevance is presumed to positively affect learners. Contemporary Expectancy-Value Theory (Eccles et al., 1983; Wigfield & Eccles, 1992, 2000) specifies four components of value that shape students' achievement-related choices and performance. Utility value refers to the valuing of a task in terms of its usefulness for helping with the fulfillment of one's immediate or future goals, and as such, closely represents the type of value that is theorized to be influenced by relevance strategies. Attainment value refers to the perceived importance of doing well on a task for one's identity. Intrinsic value refers to the enjoyment that one can get out of doing a task. Finally cost refers to the downsides of engaging in a task in terms of what might have to be sacrificed in order to do so.

While it is conceivable that teachers' emphasis on relevance could influence each of these components of value and cost, we have chosen to focus exclusively on utility value perceptions as a first step because they may be more malleable (at least in the time frame of the current study) than attainment value perceptions, which are related to one's identity. Further, relative to intrinsic value and cost, utility value seems to be more closely linked to relevance in terms of the processes involved. Thus, in the current study we focus on the relationship between teachers' use of relevance in instruction and students' perceptions of utility. Future studies might examine links between teachers' use of relevance in instruction and other components of students' value perceptions.

Much of the research within an Expectancy Value framework (and other frameworks that emphasize value) has focused on value as an independent variable rather than an outcome in itself. Indeed, Brophy (2008) writes that applications of these theoretical frameworks have "not yielded much information about how teachers might help students to appreciate the value of what they are learning" (p. 133). By linking teachers' daily practice over time with students' perceptions of utility, this study aims to fill this gap. In their investigation of relations between instructional support, task value, and student engagement, Marchand and Gutierrez (2017) argue that time-intensive investigations of the quality of classroom interactions like those employed in the current study can help illuminate the motivational processes underlying students' behaviors and beliefs in a way that contributes to both theory and practice.

Students may develop beliefs about science utility at multiple levels. Most immediate are students' beliefs about the utility of the particular course content and activities with which they are engaged on a daily basis (e.g. "why should I care which substances keep this apple from turning brown?" "How are Newton's Laws useful to me?"). Students also may develop beliefs about the utility of the domain of science on a more general level (e.g. "Can I see the way scientific understanding applies to life?"). At present, it is unclear whether teachers' use of relevance during instruction shapes students' utility beliefs at either or both of these levels. A third goal of this study is to answer this question.

1.3. Considerations in understanding links between relevance and perceived utility

While the current study examines links between teachers' use of relevance and students' perceptions of science utility, there are likely individual differences that impact both of these things. For example, students' backgrounds and self-beliefs might shape not only their everyday experience and their goals, but also the extent to which they see academic content as being useful to these experiences and goals. Likewise, teachers' perceptions of their students' capabilities, interests and backgrounds might influence whether and how teachers frame their content as relevant. Here we describe three constructs that we attempted to consider, in various ways, in our qualitative and quantitative analyses. Each of these constructs is identified in existing literature as having particular salience in discussions of value, relevance, and/or STEM learning, and thus has a potential role to play in understanding the phenomena we are examining.

Perceptions of competence. Students' beliefs about their own competence in science might influence their utility perceptions in a number of ways. Students who believe they are less competent might be less apt to recognize content as useful, either because they lack sufficient knowledge to make connections to their personal goals or because their low competence beliefs (warranted or not) lead them to believe the content is "not for me." Indeed, a number of experimental and quasi-experimental studies aimed at manipulating students' utility perceptions in a particular domain suggest that the degree to which such manipulations are effective may depend on whether or not students view themselves as competent in that domain (Durik, Shechter, Noh, Rozek & Harackiewicz, 2015; Hulleman, Godes, Hendricks & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009). In a similar fashion, teachers' beliefs about their students' competence might influence how they talk about relevance during instruction. Teachers might emphasize different kinds of relevance among students they perceive as having high vs. low competence. For example, they might emphasize applicability for careers or future education with one group while emphasizing relevance for daily living with the other.

Interest. Students do not enter their science courses in a vacuum of beliefs about the content: Rather they come with myriad beliefs and interests. It seems likely that students who are already interested in a particular task or domain might perceive that task or domain as more useful relative to others who are not interested. Scholars from a number of perspectives have posited that attempts to communicate relevance will be less influential among students who are already interested in a task and will have greatest influence on those who find the task boring or uninteresting (Deci, Eghrari, Patrick & Leone, 1994; Durik & Harackiewicz, 2007; Hidi & Renninger, 2006; Jang, 2008; Reeve, Jang, Hardre & Omura, 2002; Shechter, Durik, Miyamoto & Harackiewicz, 2011). It seems plausible that students' interest in a domain might also impact teachers' use of relevance in instruction. If teachers perceive their students to be interested, they may attempt to support this interest by more frequently emphasizing connections between content and students' goals and experiences. Alternatively, teachers might feel compelled to emphasize relevance among less interested students in an attempt to "sell" the content to them.

Representation in STEM fields. Given the longstanding underrepresentation of women in many STEM fields (see Hill, Corbett & St Rose, 2010 for a review), it is appropriate to consider whether students' gender plays a role in how teachers talk about relevance and/or in how students perceive and interpret teacher comments about relevance. For example, in prior work we found that high school science teachers tended to spend more time talking with their male students than their female students and saw their male students as more "naturally inclined" toward science (Shumow & Schmidt, 2014). Moreover, women and girls are less likely to see themselves as scientists (Miller, Eagly, & Linn, 2015; Thomas, 2017; Wang & Degol, 2017), and thus may see science as less connected to their goals.

1.4. The present study and research questions

In this study we use mixed methods to answer the following research questions:

1. How do the science teachers in the study describe the relevance of their course content and communicate these views in their daily interactions with students?
2. To what degree do the students in the study perceive their daily course content and the domain of science as useful?
3. To what extent is the way that teachers describe and communicate relevance related to students' beliefs about the utility of daily course content and the domain of science more generally?

As we endeavor to answer each of these questions, we will consider the role that perceptions of competence, interest, and gender, play with respect to teachers' relevance-related beliefs and behaviors, students' utility perceptions, and the relationship between the two.

The use of mixed methods is critical to addressing these questions. Qualitative methods enable us to gain a rich descriptive understanding of how a small number of teachers think about the relevance of their course content and how they communicate this relevance to their students during instruction. But these data alone cannot illuminate whether and how teachers' statements shape their students' perceptions. Quantitative methods allow us to gather data on the utility perceptions of a large number of students in these teachers' classrooms. The mixing of methods allows us to consider the relations between teachers' talk and students' perceptions, while also examining whether these associations vary systematically by student characteristics like prior interest, perceived competence, and gender. Mixing also facilitates triangulation, complementarity (Greene, Caracelli, & Graham, 1989) and completeness (Bryman, 2006) across multiple sources of data to provide a more comprehensive and coherent account of how teachers think about and enact relevance and how this influences students' perceptions of science utility.

2. Methods

2.1. Study design

We employed a convergent mixed methods design in which qualitative and quantitative data are gathered concurrently (Creswell & Plano Clark, 2018). The two strands are first analyzed independently, then are integrated and analyzed together. In the current study, we

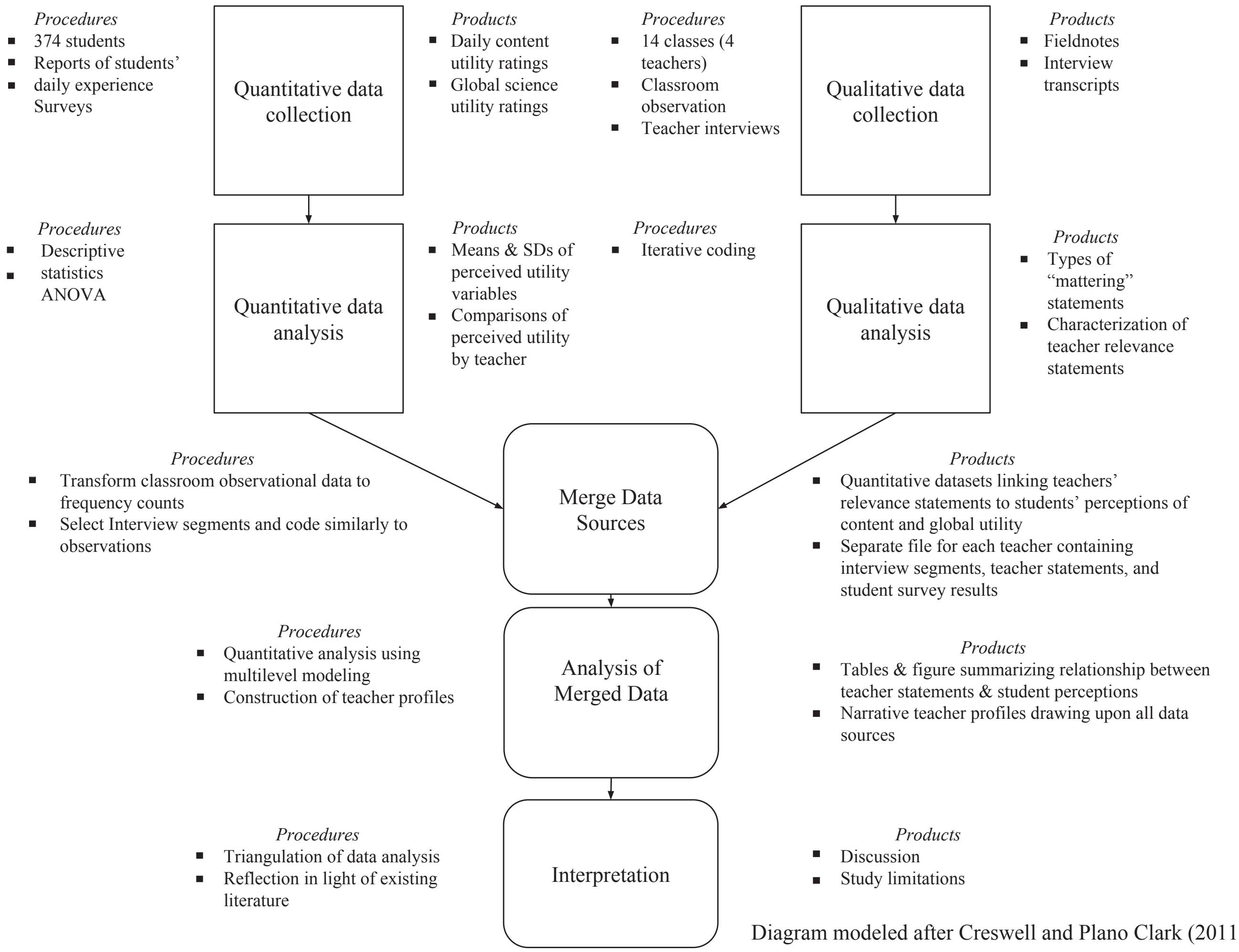


Fig. 1. Flow diagram of study design (see above-mentioned references for further information).

gathered qualitative data about teachers' expression of the relevance of science through semi-structured interviews and repeated classroom observations. At the same time, we gathered quantitative ratings of students' perceived science utility through very brief surveys that students completed on days their teachers were observed, as well as through more traditional global surveys administered at the beginning and the end of the study. The separate data strands were merged for analysis in two different ways. First, qualitative data from classroom observations were transformed into a quantitative format and were merged with the rating scale data from student surveys to build statistical models examining associations between teachers' talk about relevance and students' perceptions of science utility. Second, data from teacher interviews, classroom observations, and student surveys were used to construct integrated teacher profiles to explore how these associations unfold for specific teachers. Both of these integration strategies represent what Greene (2007) has referred to as an *interactive* level of interaction between the qualitative and quantitative data strands, meaning that there is direct interaction between them (as compared to keeping the data strands separate throughout the design). A flow diagram depicting the design is presented in Fig. 1.

2.2. Context and setting

The study took place in two middle schools from the same large, diverse school district on the fringe of a metropolitan area. In this district, science education formally begins in the third grade but students don't have specialized science teachers until sixth grade. Each school served students in grades 6–8 with enrollments near 700. The schools were demographically similar to one another, with over half of the student population identifying as Hispanic, and around two-thirds identified as low income as indicated by qualification for free or reduced-price lunch. The graduation rate of the high school into which these schools feed is 75%.

The survey, observational and interview data used in this study were gathered in seventh grade science classrooms over a seven-week period. The current investigation represents one component of a larger mixed methods study aimed at understanding sources of students' perceived utility value regarding science and sources of their beliefs about their abilities in science. The current study was conducted in conjunction with an exploration of a number of student interventions (each lasting 6 weeks), which are described in detail in Appendix A. None of the interventions was designed to influence teachers' beliefs or practices regarding science relevance (the main focus of this investigation) but they could have influenced students' utility perceptions. In consideration of both of these possible influences, the intervention group to which students were assigned was explored as an explanatory factor throughout the analysis and interpretation of both qualitative and quantitative data. The quantitative analysis showed no effects of any of the intervention conditions on students' utility perceptions (also reported in Durik, Schwartz, Schmidt & Shumow, 2018), or on the frequency with which teachers made relevance statements during instruction (see Appendix A).

2.3. Researcher positionality

As a research team, our chosen position in relation to the participants and the research context was one of outsider with substantial knowledge of the particular school context. All five co-authors of this paper identify as non-Hispanic White, are highly educated, and can be considered middle- or upper-middle class both by their upbringing and by their status as adults. As such, our lived experiences (in educational settings and otherwise) and the privileges afforded us by society undoubtedly differ from the students in our study, who are largely Hispanic and/or have relatively fewer economic resources at their disposal. Our cultural and economic backgrounds may color and/or limit the way we understand relevance and value in the context of this study. In terms of sociodemographic characteristics and the unspoken societal privileges that typically accompany them, the research team more closely resembles the teachers in the study (all non-Hispanic White women) than the students. Three members of the research team were current or former teachers at the elementary and middle school levels. Two of the researchers (the study's principal investigators) had longstanding relationships with the participating school district prior to the initiation of the current study, having conducted both research and professional development in schools in the district, though we did not have relationships with the particular teachers in our study prior to the current research endeavor. Our prior relationships with the schools undoubtedly hastened the establishment of trust between the researchers and participant teachers and provided a bit of insider knowledge to the researchers, but could also have biased our interpretations. This potential for bias was mitigated to some extent by the fact that we had other researchers on the team with no familiarity with the school or district prior to the study. Three of the co-authors of this paper were directly involved in the gathering of the qualitative and quantitative data while the other two engaged only in data coding, analysis, and interpretation of results. In sum, our research team approaches this study with varied connections to the research setting and teacher participants, and notable differences in lived experience relative to the student participants.

As a team, we approached the research process from a stance of pragmatism, recognizing the utility of diverse epistemological and methodological frameworks for understanding how teachers' beliefs and practices develop and influence students' experience in the classroom. In our shared pursuit of understanding the interactive process of teaching and learning in a way that can positively impact practice, we recognize the value of attempting to both articulate teachers' idiosyncratic belief systems and identify systematic and general relationships between teachers' enactment of these beliefs and the classroom experience of a broad range of students. Thus, we draw upon the strengths of varied approaches to empirical investigation in order to understand the phenomena of interest (Onwuegbuzie & Leech, 2005).

2.4. Qualitative data strands

Participants. Four seventh-grade science teachers participated in the study. Two taught at one middle school (pseudonyms Jackie and Celia), while the other two taught at another middle school in the same district (pseudonyms Kate and Donna). These teachers were responsible for all seventh-grade science classes at their respective schools. A total of 14 course sections were observed as part of the study. Each teacher is described briefly in Table 1.

Data collection. The general methodological framework guiding the qualitative component of this study was phenomenological in that our goal was to understand a specific construct or phenomenon (relevance), and that the focus of our study was the everyday experience of individuals (Giorgi, 1985, 1994; Moustakas, 1994).[1] Data were collected via observations and interviews. We took steps to establish trustworthiness of our process using Guba's (1981) criteria (credibility, transferability, dependability and confirmability), and Shenton's (2004) review of strategies for meeting those criteria. We describe several of the steps we took throughout our description of the qualitative methods below, and provide a more comprehensive summary of these efforts in Appendix B.

Classroom observations. Each of the 14 class sections was observed for a full 50-min class period on multiple occasions over a seven-week period. Ten classrooms (taught variously by all 4 participant teachers) were observed 14 times (twice per week). Due to scheduling constraints imposed by the broader study this investigation was a part of, 4 classrooms (taught by 2 of the teachers) were observed only 8 times (once per week, with the very first week having two observations). The observation day(s) were intentionally varied from week to week to ensure that we saw a variety of routines and schedules. We observed all classrooms studying a weather unit and a physics unit. The observational protocol entailed positioning 2–3 observers at multiple locations in the classroom to record activities and teacher-student interactions (3 observers was most typical, occurring in over 70% of the cases). Multiple observers were used to establish confirmability of observations (Shenton, 2004), and either the study's first or fourth author was involved in observation of all classrooms. Observers trained extensively with classroom video from a prior study to employ a type of event sampling (Reis & Judd, 2000) in which they provided rich description of any situation in which the teacher explicitly talked about how or why science content *mattered.* This included any instance when teachers described how or why science: applied to life, was important, valuable, useful, fun, interesting, or good. We also recorded instances in which teachers explicitly communicated that science may *not* matter (science is not useful, is unimportant, bad, etc.). Researchers made a detailed accounting of each statement (verbatim, when feasible), noting who it was addressed to, and circumstances surrounding the statement. Immediately following, observers consolidated their notes to construct a single set of field notes.

Teacher interviews. Following the seven-week observation period, each teacher participated in a semi-structured interview with one of the study's principal investigators (first and fourth authors on this manuscript). Interviews were conducted to more fully understand how teachers thought about the relevance of their science content and triangulate their stated perspectives with our classroom observations. Each interview lasted approximately 50 min and was recorded and transcribed. Using open-ended items and probes, teachers were asked to talk about activities in the weather and physics units we observed, what they hoped students would learn through those units, and what strategies were most effective in fostering that learning. We asked teachers to consider whether the strategies they mentioned were more or less effective for girls vs. boys, if gender had not spontaneously come up in their original response. We also asked them to reflect about the importance of the two units we observed, and if necessary probed for perceived connections to future education, daily life, and general scientific principles. Finally, we asked teachers to identify students who might have a future in science, and to reflect upon the qualities that prepare students for this type of future, probing specifically about male and female students.

Qualitative data analysis. Teacher statements recorded in the field notes were coded in two distinct iterative steps using the NVivo

[1] We note that our application departs from pure phenomenology in that it attempts to combine a focus on lived experience through interview and observation with other sources of data using more traditional tools of empirical investigation (see Hektner, Schmidt & Csikszentmihalyi, 2007; Creswell, 2013; Ravitch & Carl, 2016).

Table 1
Description of participant teachers.

Teacher pseudonym	Age	Total years teaching	# of 7th grade sections	Educational & professional history
School 1				
Jackie	26	4	3	Bachelor's degree in elementary education with middle school integrated science endorsement
Celia	28	6	3	Bachelor's degree in elementary education with intent to teach middle school science (no formal endorsement)
School 2				
Kate	51	13	4	Bachelor's degree in chemistry, worked as a chemist then left workforce to raise her children. Became involved with science curriculum when children were school age, then returned to school to get master's degree in education
Donna	54	20	4	Began undergraduate studies in sciences, worked occasionally as a substitute teacher to pay educational expenses and "found her niche" in education. Obtained bachelor's and master's degrees in education. At time of the study was completing a combined Master's degree in the science of education and engineering

Note: All teachers were white females.

software program. Drawing upon research on instructional strategies and effective teaching (Assor, et al., 2002; Frymier and Shulman, 1995; Keller, 1987a; Keller, Woolfolk Hoy, Goetz & Frenzel, 2016; Newmann et al., 1992; Patrick, Hisley & Kempler, 2000; van Aalderen-Smeets & van der Molen, 2015), the first coding step sorted teachers' mattering statements into broad categories using an a priori frame. Pairs of researchers (including the study's second and third authors) independently coded 20% of the fieldnotes (sampled equally across all four teachers) with 91% agreement. Four broad categories of mattering statements were identified: (1) *relevance* (science has meaning, application or value beyond the instructional context); (2) *enthusiasm/positive affect* (science is interesting, fun or cool); (3) *reluctance/negative affect* (doing science is boring, difficult, unimportant, or involves some cost or sacrifice); and (4) *other mattering.* Based on results of this first coding step we narrowed our attention for a second, more emergent round of coding, which is the focus of this article.

Because teachers referenced relevance far more often than other kinds of *mattering* (reported later in qualitative results), we developed a second coding step to code teachers' relevance statements in greater depth along several dimensions using a descriptive framework that emerged from our reading of the data (Braun & Clarke, 2006). The procedures for defining relevance statements to code, and the codes that were ultimately developed are described in Table 2. Each relevance statement was coded along six dimensions, described briefly below and labeled A-F in Table 2. Statements were coded for the *referent audience*, which indicates who the teacher was talking to at the time she made the statement (e.g., whole class, individual student, small group). We coded for the *relevance target*, which refers to who science was framed as being relevant for (e.g., relevant for everyone, for certain students, for specific others). We coded for each statement's *specificity to goals*, which refers to the extent to which the teacher connected science to goals, plans or decision making. We coded the *domain of relevance* to indicate the life domain(s) for which science was framed as relevant (e.g. everyday activities, career, school, natural phenomena). We coded *topical specificity* to reflect whether the teacher's statement referred to science generally (e.g. "orthodontists need to know a lot about science" – Donna) or to a specific topic (e.g. weather, Newton's first law). Finally, we coded the *time frame for relevance*, which indicated whether science was framed as being relevant in the present or at some point in the future. Twenty percent of the field notes (sampled equally across all four teachers) were coded by pairs of independent coders (including the study's second and third authors) using the final frame, with interrater agreement ranging from 96 to 100%, depending on the coding category.

Transcripts of the four teacher interviews were coded in the process of constructing integrated teacher profiles that draw from the qualitative and quantitative strands. These efforts will be described in the section of this article focused on the integrated analytic phase.

2.5. Quantitative data strands

Participants. The quantitative data were collected from 306 seventh grade students who were enrolled in the 14 science classrooms we observed (representing 90–100% participation within each classroom). The student sample was 45% male and 55% female. Twenty-four percent self-reported as non-Hispanic White, 17% identified as African American or multi-racial (non-Hispanic), 57% identified as Hispanic, and 2% identified as Asian. Fifty-five percent were eligible to receive free or reduced lunch. Forty-three percent reported that neither of their parents had attained a college degree and an additional 31 percent did not know their parents' educational attainment. Fourteen percent said that at least one parent had graduated from college, and 12% indicated that at least one parent had earned an advanced degree. Table 3 presents the distribution of these characteristics within each of the 14 classrooms we studied.

Data collection. Quantitative data were collected from students during the same seven-week period when their classrooms were being observed.

End of Class Reports. Students participated in a variant of the Experience Sampling Method on eight of the days when observations were made in their classrooms; twice in the first week of the study, then once per week for the remaining six weeks (Csikszentmihalyi & Larson, 1987; Hektner, Schmidt, & Csikszentmihalyi, 2007; see Schweinle, Meyer & Turner, 2006 for an example). Following each lesson, students were prompted to "think about their work in class today" and then were asked: "Do you think what you learned today could be useful outside of school?" Response options included "Yes," "No," and "Not Sure." We refer to this indicator as students' *perceived utility of daily science content.* Students also rated how interesting the day's class work was on a scale from 0 (not at all) to 3 (very much). This *daily interest* measure was included in predictive models to ensure that any observed differences in students' perceived utility were not attributable to differences in interest, and that conceptual distinctions between different types of task value (utility vs. intrinsic) could be maintained. Because End of Class reports were gathered on days that classrooms were observed, we were able to link students' subjective reports to researchers' observations about instruction at that time (this is done in the integrated analysis phase). In total, 2872 End of Class Reports were gathered.

Student surveys. Students' *perceived global science utility* was measured prior to the classroom observations, and again following the 7-week data collection period with three survey questions asking about the usefulness of their science learning and its applicability to life (adapted from Hulleman et al., 2010). Because theory and research suggest that competence beliefs and interest may moderate the way individuals think about utility value (Durik, Hulleman & Harackiewicz, 2015; Hulleman et al., 2010; Hulleman & Harackiewicz, 2009), we used items from the Time 1 survey to construct measures of *initial perceived*

Table 2
Relevance coding framework used for classroom observations.

Defining the relevance event to code: Any teacher statement indicating that science or science content is useful for some purpose (by anyone) or related to one's experience. In complex relevance statements in which multiple ideas are expressed, any shift in any of the codes below signals that the event should be coded as multiple relevance statements. (rather than multiple events). Any change in codes A-F below signals a new event. Keep in mind that we are coding instances that emphasize the utility of science for something else. We are NOT coding comments that emphasize the utility of particular tools, methods, or practices for understanding science, or for doing science better

Coded dimensions

A. Referent audience. The teacher was talking to:
- Individual male student
- Individual female student
- Individual student – gender unknown
- Whole class
- Small group – all female
- Small group – all male
- Small group – mixed gender
- Small group – unknown gender composition
- Self
- Other (includes aides, research team members)
- Don't know

B. Relevance target. Science is framed as being explicitly relevant for:
- Teacher ("this is relevant to me")
- Individual male student
- Individual female student
- Individual student – gender unknown
- Whole class/general population/everyone (the default category, the general 'you/we')
- Small group – all female
- Small group – all male
- Small group – mixed gender
- Small group – unknown gender composition
- Other (e.g., 'orthodontists, need to know this')

C. Specificity to goals:
- Real life is related to science content OR relevance of content is stated without goal
- Content *impacts* real life for specific *passive* outcome (include comfort/discomfort/expectations)—EXCLUDE if statement communicates that content impacts decisions, plans, or actions
- Content is useful for some specific goal (achievement, attainment, understanding) OR career OR if content impacts real life in a way that affects decisions, plans, or actions (of anyone)

D. Domain of relevance. Connection is made between science and:
- A job or career
- Future education in science (high school, getting into college, college majors)
- Health/safety (physical or mental)
- Routine activities/events (eating, driving), relevant cultural activities (TV, pop culture)
- Understanding or explaining natural phenomenon
- Explaining, or solving problem (local or global) that is relevant to a *specific* current event, news, or historical event (includes global warming/environmental issues). OR, explaining advances in science, health, and technology in general; emphasizing usefulness for progressing as a society
- Hobby or pastime (e.g. sports)
- Bridge to understanding a concept, unit or experience in current science class
- Bridge to understanding in another class (e.g. math, geography)
- Useful for understanding or advancing social relationships

E. Topical specificity. Connection refers to:
- Specific topic in science
- Science in general ("orthodontists need to be good at science")

F. Time frame for relevance:
- Implied whenever, no time mentioned, past, or general "in the future" or "right now" suggesting all of the time
- Immediate future (this week or month)
- Moderate future (this season, year, time through high school)
- Long term (in college, career, as adult)

competence (items adapted from Wigfield & Eccles, 2000) and *initial global science interest* (adapted from Harackiewicz, Durik, Barron, Linnenbrink-Garcia & Tauer, 2008). Item descriptions, descriptive statistics, and psychometric properties (when applicable) for all quantitative measures are presented in Table 4.

Quantitative data analysis. Our initial analysis of the quantitative data consisted of generating descriptive statistics (means, standard

Table 3
Student demographic characteristics by teacher and class period.

	School 1						School 2								
	Jackie			Celia			Kate				Donna				
	P4 (n = 25)	P7 (n = 22)	P8 (n = 28)	P3 (n = 20)	P4 (n = 21)	P7 (n = 23)	P1 (n = 20)	P3 (n = 17)	P6 (n = 23)	P7 (n = 18)	P3 (n = 22)	P4 (n = 20)	P6 (n = 24)	P8 (n = 23)	**TOTAL** (N = 306)
Gender															
Male	40	54	40	40	43	44	40	48	48	61	37	30	50	48	45
Female	60	45	60	60	57	56	60	52	52	39	64	70	50	52	55
Race/ethnicity															
Non-Hisp White	12	18	14	20	24	5	45	7	35	28	18	50	33	35	24
Af-Am/mixed	20	36	11	30	28	30	5	23	9	22	18	15	21	9	17
Hispanic	60	41	68	50	48	65	50	70	52	50	64	35	46	56	57
Asian	8	5	7	0	0	0	0	0	4	0	0	0	0	0	2
Free/reduced lunch	61	45	75	44	61	72	50	50	45	53	44	58	21	52	55
Parent Educ.															
Don't know	24	23	36	30	55	30	10	26	31	17	24	40	54	35	31
< college deg.	52	41	46	40	35	53	40	68	43	45	49	30	16	39	43
4y coll. deg.	8	23	7	25	10	13	35	0	17	11	14	15	9	9	14
Advanced deg.	16	13	11	5	0	4	15	6	9	27	13	15	21	17	12

Notes: Values reported in each cell are percentages.

Table 4
Descriptions of quantitative measures used in analyses.

Measures from survey	Range	M	SD	α
Initial perceived competence[a]	1–7	4.34	1.20	.80
How good in science are you?				
If you were to list all the students in this class from the worst to the best in science, where would you put yourself?				
Compared to most of your other school subjects, how good are you in science?				
Initial global science interest[b]	1–7	4.46	1.61	.92
I'm really excited about learning science				
I've always been fascinated by science				
I think science is an important subject				
Global science utility[c] **(Initial – I & Post - P)**	1–7	4.74 (I) 4.90 (P)	1.30 (I) 1.30 (P)	.81 (I) .85 (P)
I can see how what I learn from science applies to life				
I think what we are studying in science class is useful to know				
I can apply what we are learning in science class to real life				
Repeated measures from end of class reports	**Range**	**M/freq.**	**SD**	**α**
Daily interest	0–3	1.72	1.07	–
Was the work in class today interesting?				
Daily utility				
Do you think what you learned today could be useful outside of school?	No Yes Not sure	22% 45% 33%	–	–

[a] Adapted (to refer to science) from Wigfield and Eccles (2000).
[b] Adapted (to refer to science) from Harackiewicz et al. (2008).
[c] Adapted (to refer to science) from Hulleman et al. (2010).

deviations, range) to characterize students' perceptions of the utility of their daily content and of science more generally. Using oneway ANOVA, we also tested for mean differences in these student ratings by teacher. We examined additional sources of variance in students' utility ratings in the integrated analytic phase, described below.

2.6. Integrated analytic phase

We conducted two kinds of analysis in the integrated phase. Using data from classroom observations and student surveys, we employed data transformation and merging to examine how science teachers' instructional talk about the relevance of science is related to students' utility beliefs. Next, we used the teacher interviews, observations and student surveys to construct integrated teacher profiles to explore the unique ways the four teachers in our study thought about and enacted relevance in science instruction.

Data Transformation: Quantizing. We converted our observational data into quantitative format so that our two strands of data could be merged and analyzed together using advanced statistical techniques. For each class session observed, we counted the number of coded relevance statements of any kind, and then matched each count to students' daily utility ratings for that class session.[2] We used Hierarchical Linear Modeling (HLM; Raudenbush & Bryk, 2002) to examine whether students' daily utility perceptions were related to a number of factors including the number of relevance statements their teacher made on that particular day, who their teacher was, and individual characteristics that our literature review suggested might be important (perceived competence, interest in the task at hand, and gender). Because the daily utility ratings were dichotomous (yes vs. no/not sure), we used a three-level logistic hierarchical and cross-classified random effects model in which students' responses were cross-classified by person and day nested within classroom. Random effects for student, day, and classroom were included in the model. The level "day" refers to a specific course section on a particular day: This is the level at which teachers' relevance statements are modeled. This analytic structure models the non-independence of observations in two different ways. While multiple observations coming from the same person are not independent, neither are multiple observations coming from different students in the same classroom on the same day (i.e., everyone in Celia's third hour class on Tuesday experienced the same relevance statements). Since each teacher had multiple science sections in the study, teacher effects can be examined across classrooms.[3]

To merge the daily relevance counts with students' global utility ratings (measured at the beginning and end of the study), we calculated the average daily relevance count by classroom across all observation days. The merged data were analyzed using a two-level hierarchical linear model predicting the continuous Time 2 measure of students' global science utility. Level 1 (person) predictors included the student's initial science utility rating and the individual characteristics identified in our literature review (perceived competence, interest, and gender). Level 2 (classroom) predictors included average daily relevance statements and teacher.

A few words are warranted about how we understand the observational counts used in the integrated analyses. In the context of this study, the only way to meaningfully interpret the number of statements a teacher makes in a given class session, or in a particular course over time, is relative to the number of statements she (or another teacher) makes in a different class session, or in another course over time. The values themselves have no absolute meaning beyond signaling "relatively more" or "relatively less." We do not aspire to identify an "optimal" or "threshold" number of statements for achieving effects on students' utility, and do not believe such an endeavor would yield any useful practical application.

Construction of teacher profiles. Following procedures outlined by Seidman (2013), our final analytic step was to construct profiles of each of the four teachers representing their thinking about and expression of relevance using data from the teacher interviews, classroom observations, and student surveys. Profiles facilitate coherent presentation of the participants in context, clarify participants' intentions, and convey a sense of process (Seidman, 2013).[4] The teacher interviews served as the foundation for the profiles, and then the other data sources were integrated to construct comprehensive and nuanced accounts of teacher beliefs and practices and how they affect students. Transcripts of the four teacher interviews were coded by two researchers who had not been involved in the coding of the observational data (to avoid potential bias). We first read the transcripts in their entirety to familiarize ourselves with the data and then highlighted passages in which teachers discussed: (a) strategies they used to promote student learning, (b) why

[2] Note that some classrooms were observed on days in which End of Class Reports were not collected. While these additional observation days informed our qualitative analysis, any observational counts that could not be matched to data from End of Class Reports were not included in this merged analysis.

[3] We remind the reader here that we also examined whether teacher talk varied systematically by the intervention condition to which classrooms were assigned as part of the larger study. Teacher talk was associated with teacher rather than intervention condition (each teacher was assigned to multiple conditions), so the teacher indicator was chosen for inclusion in analysis.

[4] Seidman recommends multiple in-depth interviews to support the construction of profiles and uses the term "vignette" to refer to profiles that represent a narrower range of participants' experiences and/or are constructed from shorter narrative data. Thus, our profiles more closely represent what Seidman refers to as "vignettes." Because in educational psychology research the term vignette is more often used to refer to fictional stimulus materials used in quantitative research we use the term "profile" here to avoid confusion. Some researchers choose to present profiles as a first-person narrative but we have chosen to retain the third person voice throughout this article and represent participants' own words using quotation marks.

Table 5
Total number of observed "mattering" statements and average statements per class period, by teacher.

Teacher	Relevance N (avg. #/class)	Enthusiasm/Pos. affect N (avg. #/class)	Reluctance/Neg. affect N (avg. #/class)	Other N (avg. #/class)	Total N (avg. #/class)
Jackie[a]	24 (0.7)	5 (0.2)	2 (0.1)	4 (0.1)	35 (1.0)
Celia[b]	47 (1.7)	1 (0.04)	15 (0.5)	7 (0.3)	70 (2.5)
Kate[c]	38 (0.8)	15 (0.3)	9 (0.2)	8 (0.2)	70 (1.5)
Donna[d]	183 (4.6)	18 (0.5)	10 (0.3)	8 (0.2)	219 (5.5)
Total[e]	292 (1.9)	39 (0.3)	36 (0.2)	27 (0.2)	394 (2.6)

Notes: This table reports the total number of units coded for each teacher. Because teachers taught a different number of class periods and were observed a different number of times, this number is not a good indicator of how often a given teacher made statements in a given class period. The number reported in parentheses is the average number of statements observed per class session.
[a] Jackie was observed in a total of 36 class sessions (3 course sections observed 12 times each).
[b] Celia was observed in a total of 28 class sessions (1 course section observed 12 times and 2 course sections observed 8 times each).
[c] Kate was observed in a total of 48 class sessions (4 course sections observed 12 times each).
[d] Donna was observed in a total of 40 class sessions (2 course sections observed 12 times each and 2 course sections observed 8 times each).
[e] The total number of class sessions observed across all teachers was 152.

it was important for students to learn the content they were presently studying, (c) qualities of specific students that they perceived as having a future in science; and (d) gender as it relates to students' interest, ability, or aptitude in science, or regarding students' responsiveness to particular instructional strategies. Note that our interview protocol included direct questions about each of these topics, but that these topics also came up at other points during the interview. This focused approach is appropriate in cases where previous theory and research was used to shape the research questions of a study that involves qualitative methods (see Hsieh & Shannon, 2005). Passages related to these four topics (n = 53) were identified through consensus and were assembled in a separate document for each teacher. Passages were then coded by one of the researchers using a frame that was derived from our observational coding protocol, as our questions about why the units we observed were important closely mirrored the kinds of statements we coded in our classroom observations (see Table 2). New codes were generated by the coder to characterize teachers' reflections about specific teaching strategies, about gender, and about students with futures in science because these topics were not readily addressed in the observational protocol. Once the final coding frame had been established, it was applied to all passages by the second researcher. The two independent coders were in agreement about how to code 91% of the passages. The remaining 9% were resolved through discussion.

At this point the other data strands were integrated into the profile documents. Results of classroom observation coding were summarized both within and across each teachers' classrooms, which allowed us to examine the extent to which each teacher's practice varied across her different course sections. For each teacher, we summarized students' ratings of perceived daily utility and global science utility, and then also included brief descriptions of the students in each teachers' classrooms in terms of demographic characteristics and number of students characterized as having "gifted" or "special education" status.

This corpus of data was then reviewed by the two researchers who had coded the interview data to look for patterns, consistencies, contradictions, and questions that arose across the multiple data sources. The result was the identification of four broad themes around which to organize the final profile narratives. It is typical as a final step in the construction of profiles to generate broad themes as a way to bring meaning, structure and coherence across multiple profiles (Seidman, 2013). The four organizational themes included: (1) effective instructional strategies; (2) why science content is important; (3) who science is for (and not for); and (4) gendered experiences in science. The generated themes were reviewed in light of all of the coded extracts, and then the full interview transcripts were read again in their entirety to check that nothing had been missed or mischaracterized (see Braun & Clarke, 2006; Seidman, 2013). This final check resulted in the selection of six additional interview segments for coding and inclusion in construction of the profile narratives.

3. Results

3.1. Qualitative results: How teachers used relevance in instruction

Using the fieldnotes from our classroom observations, we identified and coded 394 unique instances when teachers made statements about why science mattered. This amounts to an average of 2.6 observed statements in each class period we observed. Two hundred ninety-two of these statements (or 74%) were categorized as relevance statements because they made an explicit connection between science content and students' experiences or goals outside of the immediate learning context. Thirty-nine statements (10%) were coded as expressing teacher enthusiasm or positive affect, as they explicitly articulated that science was fun or interesting. Thirty-six statements (9%) expressed reluctance or negative affect about science or emphasized the costs of engaging with the content. Most often teachers told students that science was extremely difficult, would take time away from other things they might want to do, or would "stress them out." We identified 27 instances (7%) in which teachers clearly said science content was important, but we were unable to fit their statement into any meaningful category. In most of these statements the teacher said something akin to "this stuff is really important" without explaining why. Table 5 shows the types of mattering statements made by each teacher.

Because teacher statements about relevance were made more frequently than other kinds of *mattering* statements, we developed a more detailed coding frame to characterize the nature of those statements. Results of that process were used to address our first research question about how teachers describe and communicate relevance. Most of the relevance statements teachers made (77%) were directed at the entire class rather than individual students. In those rare occasions when teachers communicated relevance directly to an individual student, these statements were nearly always made publicly in front of the whole class, so everyone in class heard the message being conveyed. Teachers rarely made relevance statements in which they characterized science as being relevant for a specific person (either a student or another identifiable individual). Rather, in 90 percent of the instances we recorded, teachers' relevance statements were more general, suggesting relevance for anyone. Because teachers' relevance statements so rarely identified relevance in terms of specific persons, there was no evidence that teachers' expressed relevance statements inequitably by individual student gender.

Teachers' relevance statements nearly always (93%) referenced specific science content (e.g. "Newton's laws help you understand what happens to the ball when you play soccer." - Donna), rather than referring to the relevance of science more generally (e.g., "Don't they know that orthodontists need to know about science?" - Donna). Most of the statements teachers made about relevance (82%) either explicitly referred to the present or did not refer to a specific time frame (e.g. "It's like when you slam on the brakes really hard in the car – what happens to the passengers inside?" – Kate). Only 18 percent of all the statements we observed referred specifically to relevance in the future (e.g. "you'll need to know this for high school" – Celia).

Teachers' relevance statements referenced connections to routine activities and hobbies like sports and driving a car (24%); natural phenomena and local or global problems like climate change (19%); health and safety concerns (16%); specific jobs or careers (11%), and current or future education (7%). About 23% of all teacher statements were not readily identifiable as pertaining to a specific domain of life but were nonetheless referencing relevance (e.g. "I wish I could get you guys to understand that speed is about more than cars" - Donna).

Teachers made statements with varying levels of specificity regarding the connection of science content to goals, plans or decisions. We coded each relevance statement as falling into one of the following levels of specificity, with about one-third of all relevance statements falling into each category:

- *Low specificity*: Teachers made a connection between science content and students' experience, but with no reference to the impact of science on any particular event, goal or outcome (e.g. "Like yesterday when you were going home from school – did you see them? What kinds of clouds?" student responds "stratus." "Yes!" - Celia).
- *Moderate specificity*: teachers describe how science impacts the self for a specific passive outcome but make no reference to the use of this knowledge for making plans or decisions (e.g. "high humidity can make your hair frizzy" –Jackie). While these statements identify a cause-effect relationship involving science constructs, there is no acknowledgement of how it might relate to one's plans, goals, actions or decisions.
- *High specificity*: Teachers describe how science content is related to specific goals, decisions, plans or actions (e.g. "if you know what cumulonimbus clouds and funnel clouds look like you might be able to protect your family from tornadoes and other bad storms." – Celia). Such statements explicitly state a situation which scientific understanding could be useful.

In comparing relevance statements across the four teachers in the study, the most striking pattern is that Donna talked about relevance far more often than the other teachers. On average, Donna referenced relevance four or five times in each of her 50-min class sessions. In contrast, Jackie and Kate referenced relevance less than once per class period, and Celia averaged between one and two references per class. We dive deeper into teacher-related differences regarding relevance processes in our quantitative analysis of student survey data (presented next) and then again through the construction of comprehensive profiles that integrate teachers' interview responses with our classroom observations and quantitative results.

3.2. Quantitative results: students' ratings of science utility

Quantitative analyses of data from the student surveys and End of Class reports were used to address our second research question about the degree to which students perceived their daily course content and the domain of science as useful.

Students' daily science utility ratings. Looking across the nearly 3000 End of Class Reports we collected from multiple students in multiple classrooms, 46% indicated "yes," that the day's content could be useful outside of school, while 22% indicated "no" and 32% indicated "not sure." For the quantitative and merged analyses these response categories were collapsed to create a dichotomous indicator of daily utility where "yes" is equal to 1 and "no" and "not sure" are equal to 0. This decision was made on both conceptual and empirical grounds. Conceptually, the "yes" response is the only one that clearly indicates perceived utility. Also, preliminary analyses examining all three available categories indicated that the "no" and "not sure" response categories showed similar relations to other variables of interest, so treating these response categories separately would have complicated analyses considerably without adding to our understanding.

Even though all classrooms were covering the same content when data were collected, simple means comparison reveals that the extent to which students perceived their daily content as useful was systematically related to which teacher they had (ANOVA, $F = 75.48$, $p < .001$, see Fig. 2). Of all the End of Class Reports collected in Donna's classrooms, 63% indicated that students saw what they were doing as useful, whereas daily utility was indicated in only 49% of responses from Celia's classes, and 32% of responses from both Jackie's and Kate's classes (post-hoc comparison indicated that all teachers differed significantly from one another with the exception of Jackie and Kate). We also explored the extent to which students' daily utility ratings were related to a variety of other situational, individual and classroom factors. These results are reported in detail in the integrated analysis section below.

Students' global science utility ratings. At the outset of the study, students reported an average rating of 4.74 on the three-item indicator of global science utility (SD = 1.34, range 1–7). As the midpoint of the scale, a value of four indicates relative neutrality, so the sample mean of 4.74 indicates that on average students agree slightly with statements asserting that the domain of science is generally useful and/or applicable to life.

Students' global utility ratings were also related to which teacher they had ($F = 7.9$, $p < .001$, see Fig. 3). Simple means comparisons indicate that Donna's students reported significantly higher utility ratings ($M = 5.33$) than Jackie's ($M = 4.47$) or Kate's ($M = 4.5$). Celia's mean student ratings fell in between these two extremes ($M = 4.76$) but was not significantly different from any of them. We also explored the extent to which students' global utility ratings were related to a variety of other individual and classroom factors. These results are reported in detail in the integrated results section below.

3.3. Results of the integrated analyses

The integrated analyses were aimed at addressing our third research question regarding whether and how teachers' attempts to communicate their beliefs about relevance were related to their students' utility beliefs.

Teacher statements and students' daily utility. Table 6 displays the results of a three-level logistic and cross-classified model predicting students' daily utility ratings as a function of a number of situational, personal, and classroom characteristics. The log odds coefficients from the model were converted into odds ratios, which can be interpreted as representing increases or decreases in the likelihood of finding the day's work useful. On any given day, if a teacher made more relevance statements during the natural course of instruction, the students in her class were more likely to view that day's material as being useful outside of school ($\delta_{040} = 1.11$, $p < .05$). For every additional relevance statement a teacher made, the odds of students reporting that the day's lesson is useful increased by 11%. This positive association is statistically significant even after controlling for teacher, level of student interest during that particular day's lesson, and students' beliefs about their own science competence and the utility of science more generally.

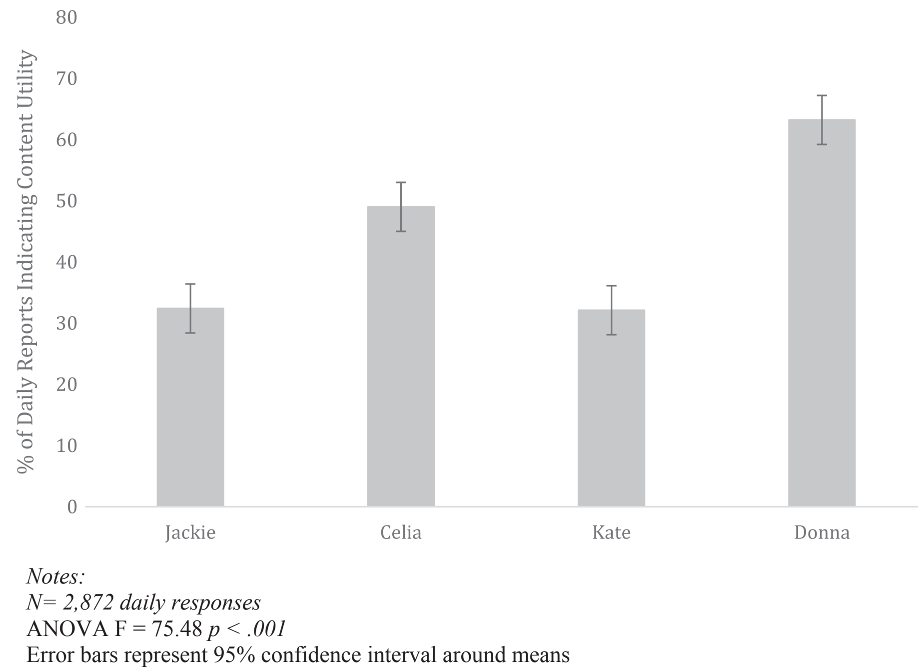


Fig. 2. Proportion of students' daily reports indicating content was useful, by teacher. Notes: N = 2,872 daily responses ANOVA F = 75.48*p* < *.001* Error bars represent 95% confidence interval around means.

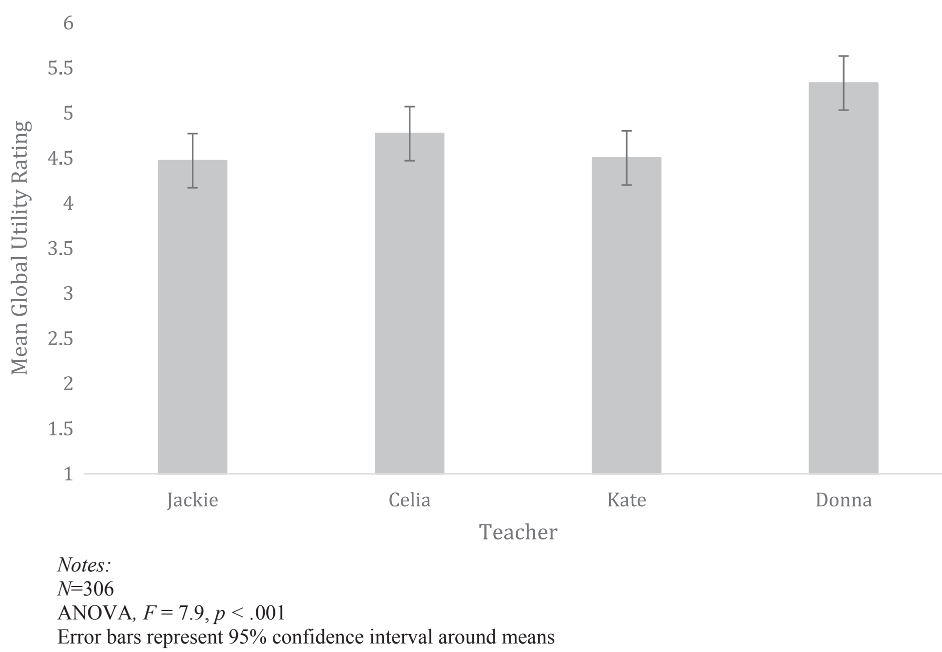


Fig. 3. Students' mean initial global utility ratings, by teacher. *Notes: N* = 306 ANOVA*, F* = 7.9, *p* < .001 Error bars represent 95% confidence interval around means.

Even when the number of teacher relevance statements is accounted for, students in Donna's (δ_{003} = 2.66, $p < .05$) classrooms were still more likely to see their class content as useful relative to Jackie's students, which suggests that Donna may have used additional strategies beyond those we defined as relevance that also had positive effects on students' utility beliefs. Students' daily utility ratings were unrelated to their gender (δ_{010} = 0.91, *ns*). Not surprisingly, students who began the study with higher levels of initial global science utility were more likely to see their daily content as useful, net of the number of relevance statements their teacher made (δ_{030} = 1.72, $p < .001$). Additionally, when students reported being more interested in a particular day's content, they were also more likely to see that content as useful, net of the number of relevance statements a teacher made (δ_{100} = 1.80, $p < .001$). The general association between students' perceived competence and their daily science utility was not statistically significant, but the variance component for that relationship was (d_{01} = .03, $p < .05$), indicating that the relationship between competence and daily utility perceptions differs across students.

Teacher statements and students' global science utility. Table 7 presents a two-level HLM analysis in which mean number of daily relevance statements over the seven-week observation period was used as a classroom-level (level 2) predictor of students' global utility ratings at

Table 6
Associations of teachers' daily relevance statements with students' perception of the utility of that day's science content.

Fixed effects	Odds ratio of finding the day's work useful
For Intercept1	
For Intercept	
$\theta 0, \delta_{000}$	0.51*
Celia,[a] δ_{001}	1.81
Kate,[a] δ_{002}	1.10
Donna,[a] δ_{003}	2.66*
For female slope	
Intercept, δ_{010}	0.91
For perceived competence slope	
Intercept, δ_{020}	1.00
For initial global science utility slope	
Intercept, δ_{030}	1.72***
For teacher's daily Relev. Stmts. slope	
Intercept, δ_{040}	1.11*
For daily interest slope	
For intercept	
$\theta 1$, δ_{100}	1.80***
Random effects	*Variance components*[c]
Within students by day	_[b]
Between days, θ_0,c00	0.03*
Between students, θ_0,b00	1.80***
Between classrooms, θ_0,d00	0.02**
Between students, competence, d_{01}	0.03*

Notes: Estimates are for a logistic (logit link) three-level hierarchical and cross-classified random effects model, with students' repeated responses nested within cells crossed by both student and day, and day nested within classroom; The set of teacher indicators (Celia, Kate, Donna) are classroom level variables, as each teacher taught multiple classrooms and there were too few teachers to model at a higher level. The variables female, perceived competence, and initial global science utility are student level variables. While the variable Hispanic was included in the initial models, it was removed from the merged analysis because it did not explain any variance in daily utility in any models. Daily interest is modeled at the level formed by crossing student with day. Students' perceived utility of daily science content was measured using students' response to the question "Do you think what you learned in class today could be useful outside of school?" The variable is coded 0 = no/I don't know, 1 = yes. The independent variables competence, initial utility, teachers' relevance statements, and interest were centered around the grand mean. For ease of interpretation, all HLM coefficients were converted to the odds ratio of a student indicating that the day's content could be useful outside of school (relative to not useful or I don't know).

[a] Reference teacher is Jackie

[b] In logistic models, this variance component is not estimated, but is fixed at $\pi^2/3$.

[c] Intra-class correlations calculated using variance components from the unconditional model indicated that about half the total variance occurred within students (3.29), over one third occurred between students (2.48), while variance in between classes (0.58) and between days (0.06) accounted for 9.07% and 0.91% of the variance respectively. Because the level-1 variance is fixed at $\pi^2/3$ for any logistic model, determination of variance reduction due to addition of covariates and random effects was not calculated.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

the end of the study. Because initial global utility ratings were included as a predictor in the model, the coefficients presented in Table 7 can be interpreted as representing change in perceived global science utility from the beginning to the end of the study. As expected, students' initial utility perceptions were predictive of utility at the end of the study ($\gamma_{40} = 0.45, p < .001$). Students who had higher science interest at the outset of the study showed larger increases in global science utility over time ($\gamma_{30} = 0.12, p < .05$). Classroom teacher was not associated with change in global utility beliefs over time once the variation due to teachers' relevance statements was taken into account. As was the case

Table 7
Associations of teachers' average daily relevance statements with changes in global science utility over a 7-week period.

Fixed effects	B
For Intercept1, β_0	
Intercept, γ_{00}	4.79***
Celia,[a] γ_{01}	0.30
Kate,[a] γ_{02}	0.04
Donna,[a] γ_{03}	0.44
Teacher's Avg. Daily Relev. Stmts., γ_{04}	−0.24
For female slope, β_1	
Intercept, γ_{10}	−0.20
Teacher's Avg. Daily Relev. Stmts., γ_{11}	0.28*
For perceived competence slope, β_2	
Intercept, γ_{20}	0.06
For global science interest slope, β_3	
Intercept, γ_{30}	0.12*
For initial global science utility slope, β_4	
Intercept, γ_{40}	0.45***
Random effects	*Variance components*[b]
Between classrooms, u_0	0.04*
Between classrooms, competence, u_2	0.03*
Within classrooms, *r*	1.15

Notes: The results presented in this table are an expansion of analysis originally presented in the dissertation of Stephen Kafkas (2016), for which Jennifer Schmidt served as dissertation advisor. Estimates are for a two-level Hierarchical Linear Model with students at Level 1 and classrooms at level 2. The set of teacher indicators (Celia, Kate, Donna) are classroom level variables, as each teacher taught multiple classrooms and there were too few teachers to model at a higher level. The variables female, perceived competence and global science interest are student level variables. The independent variables competence and interest were centered around the grand mean. While the variable Hispanic was included in initial analysis, it was removed from the merged analysis because it did not explain any variance in daily utility in any models.

[a] Reference teacher is Jackie.

[b] Intra-class correlations calculated using variance components from the unconditional model indicated that most of the variance occurred within classes (1.79) and about 6% occurred between classes (0.12). Addition of covariates and a random effect for competence reduces total within-class variance by about 36% and between-class variance by 81%.

* $p < .05$.

*** $p < .00$.

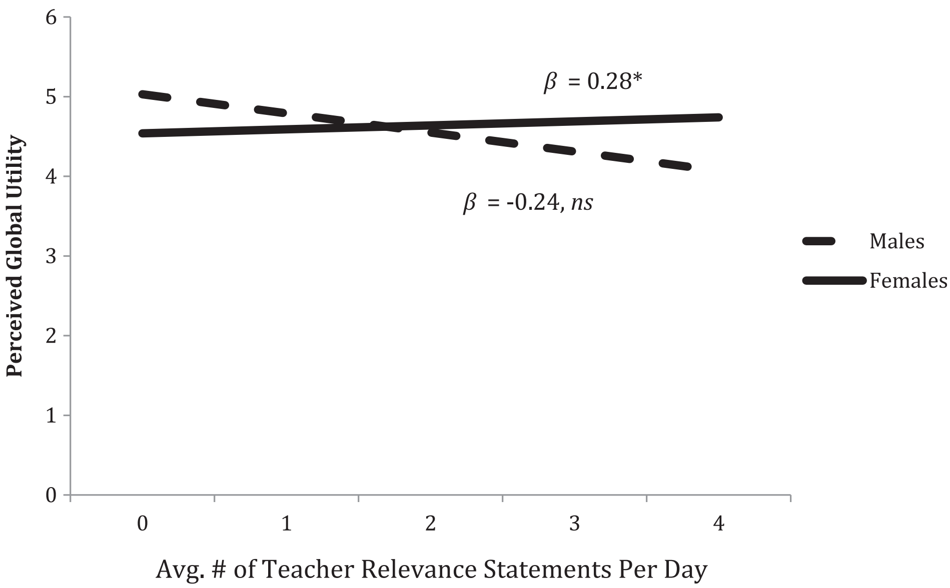
in our analysis of daily utility, students' competence beliefs were generally not associated with global utility beliefs, but the variance component for this association was statistically significant ($u_3 = 0.03$, $p < .05$), suggesting that the relationship between students' perceptions of competence and global science utility differs across classrooms.[5]

Due to the way the models in Table 7 are specified, the non-significant coefficient for the intercept γ_{04} shows that there is no discernible relationship between teachers' relevance statements and change in science utility beliefs for male students. For female students, however, teachers' average number of daily relevance statements was associated with significant increases in global science utility beliefs over time ($\gamma_{11} = 0.28, p < .05$). Fig. 4 represents the interactive effect of mean number of daily teacher statements and gender on students' perception of global science utility.

Integrated teacher profiles. We present profiles for each teacher below.

Jackie. Jackie's interview reflection about the importance of her science content exclusively emphasized fairly simple, practical implications for students: "Knowing about weather tells you how to dress." She expressed the belief that connecting science lessons to

[5] We remind the reader again here that the classroom interventions that were being conducted alongside this study were not associated with changes in global science utility.



Notes: Dependent variable (perceived science utility) measured at the end of 7 week observation period. Analyses control for teacher, as well as students' perceived global science utility, science interest, and perceived science competence at the outset of the study
* p < .05

Fig. 4. Effects of teachers' relevance statements made during instruction on students' perceptions of global science utility by gender. *Notes:* Dependent variable (perceived science utility) measured at the end of 7 week observation period. Analyses control for teacher, as well as students' perceived global science utility, science interest, and perceived science competence at the outset of the study.

students' everyday lives was motivating and helped them learn better. She described with enthusiasm how throughout her weather unit she had students track their local weather using The Weather Channel's website and local newspapers. Jackie contended that her students were more enthusiastic and learned more (as indicated by test scores, she said) when they got to track something that was happening right outside of their window. We observed this activity on multiple occasions in all of Jackie's course sections and in our fieldnotes recorded little to no discussion of what the forecast information meant or how it should be used. Students in all her course sections simply recorded each day's high and low temperatures and precipitation in a journal and, by all indications the exercise appeared to elicit the same level of student engagement as roll call or the daily collection of completed homework.

Jackie struggled to see (or at least to articulate) how her content fit into students' educational trajectories or how it contributed to knowledge of scientific principles more generally. She reflected that students' understanding of weather would apply to a unit on plants that would occur in spring of that same year but did not see any connection to other work students would do in middle school or high school. Jackie appeared unsure about how knowledge about weather might be important for understanding general scientific principles or for building a repertoire of scientific knowledge, stating in a rather questioning tone that the weather unit could be useful for "maybe patterns?"

The portrait of a fairly narrow and shallow view of the relevance of one's subject matter that came out of Jackie's interview is largely supported by our classroom observations. Of all the teachers in the study, Jackie made the fewest relevance statements – only 24 in total. The number of relevance statements she made in a given class period ranged between 0 and 3, but across all of these class periods she averaged 0.7 relevance statements per class period, with her most frequently observed "relevance count" for a class period being 0 (see Table 5). In fact, we observed Jackie's 8th period science class on 14 different occasions over a seven-week period and did not observe a single relevance statement. While Jackie generally made far fewer relevance statements than her peers, a relatively large proportion of her statements had high specificity in their connection to goals, plans and decision-making (60% of her relevance statements were coded this way, compared to 8–38% for other teachers). For example, during her unit on weather Jackie stated "a lot of people look at the weather for their professions or recreation," and proceeded to explain the weather-related factors that the school's soccer coach must weigh when making decisions about whether or not the team can play when it is raining. Jackie almost never made relevance statements that were coded as having low specificity of utility, meaning that she rarely pointed out examples of science content in students' daily experience without reference to specific actions, plans or outcomes. Jackie's tendency to focus on a more limited range of applications seems consistent with the narrow range of connections Jackie was able to articulate in her interview.

Jackie's relatively narrow views about the importance of her science content and limited use of relevance as an instructional strategy appear to be reflected in her students' beliefs about the utility value of science. Overall, students in Jackie's classes reported among the lowest levels of daily science utility (32% of all responses, tied with Kate's students for lowest ratings, see also results presented in Fig. 2 and Table 5). To put this in context, Jackie's students reported that their content was useful half as often as Donna's (63%), and substantially less often than Celia's (49%). Similarly, Jackie's students had the lowest average global science utility ratings (M = 4.5, again tied with Kate), which were lower than Donna's students (M = 5.3) and Celia's (M = 4.7), though the difference from Celia was not statistically significant.

More than any of the other teachers, Jackie's interview responses suggest a belief that a fairly limited group of students has potential for a future in science. Jackie felt that the school's classification of students as "gifted" in either language arts or mathematics (a status that was assigned relatively generously, judging by the numbers of students who qualified as gifted) was likely the best indicator of which students could have a future career in science. In her interview she described how she differentiated student activities and assessments by gifted status (though she did not use the term 'differentiate' – she simply said she gave the gifted students "modified" tests and "extra" assignments). We frequently observed this practice in her classrooms, and she was the only teacher we saw doing this. She singled out gifted and non-gifted students in very public ways – calling them by name, asking students to remind her if they were gifted, physically separating them, and giving gifted students assignments that the other students perceived as more fun (e.g., creating colorful brochures about climate vs. vocabulary worksheets). Jackie's reflection about which of her students may have a future in science also suggested that she may embrace some stereotypes about social class in that she added that parents were another important influence on whether students would have a future in science. The examples she provided were forms of involvement and enrichment that are prevalent among middle class but not among working- or lower-class parents (Lareau, 2011).

Given the salience for Jackie of students' status as gifted, it is noteworthy that her 8th period class – the one in which she made 0 relevance statements in 14 class sessions over a period of seven weeks – had only 4 gifted students. Her other two course sections had 16 and 13 gifted students, respectively, and in those classrooms, Jackie made an average of 1 relevance statement per day over the same seven-week period. If indeed Jackie believes that gifted students are best suited for science, this belief may have (consciously or not) impacted how she framed her science content such that she simply did not discuss the relevance of science with the group of students she viewed as least likely to have a future in that domain.

Though Jackie mentioned both male and female students in our discussion about who might have a future in science, when asked explicitly about gender differences in students' science experience she described how a girl in her 8th period class "asked me why did we need to know science… and why it was important." She went on to explain that she thought this view was particularly prevalent among female students across her classes though most would not be so bold as to mention it, and that in her experience boys enjoyed science more and asked more questions. When asked about whether she does anything to address these perceived differences, she described her solution in terms of emphasizing relevance:

> *I always tried to tell them that science is everything that they do, it's them eating, them breathing, them, you know, going outside, you know, the building that we're in, um, so I tried to get them to understand that it's not just what we're learning in the classroom, it's them… Um, and so you know I always try to bring back a real-world example, um, 'cause sometimes when you're talking in scientific terms they might just get lost… and not feel that I'm talking to them.*

Of course, our merged quantitative analysis provides some support for the effectiveness of this particular strategy: Making more relevance statements increases daily utility perceptions for all students and raises global utility perceptions among girls in particular. However, our observations of Jackie's classroom suggest that in practice, she employed this strategy only rarely.

Celia. When asked to reflect on strategies she thought were effective for facilitating learning in the weather unit, Celia's response was:

> *Getting them to make it personal. We went outside a lot, we went, um, when we did research on other towns or states, I encouraged them to pick states that they knew about, like they had family members or um, previously lived. So, I definitely tried to make everything personal to them… that's the best way [of helping students learn the content]… making it personal.*

Celia's belief was that "making it personal" was generally effective because it made students more interested in doing course activities and resulted in increased understanding. While her spontaneous response to this question did not reference gender, in response to a probe Celia observed that her "making it personal" strategy tended to have a greater positive effect on boys than girls. In her experience boys were much more likely to complete homework tasks if they emphasized relevance to students' lives, but girls did not seem to require that connection —they just did their homework regardless of whether or not they perceived it as relevant. While on the surface this statement indicates that she views girls as being more conscientious students than boys, it also implies a belief that girls may not "need" to hear about science relevance as much as boys do.

Like Jackie, in her interview Celia articulated the simple practical importance of understanding weather, noting its importance "so students can decide whether they need an umbrella while walking to the bus." Beyond these simple daily applications, Celia also suggested that knowledge of science might better equip students to take more informed positions on larger global issues. She shared the belief that it was valuable for students to study climate change (which she incorporated into the weather unit we observed – she was the only teacher who did this) because this knowledge would help students "decide whether they believe it or not." Celia expressed fairly limited connections to students' future education or general scientific principles. She noted, as Jackie did, that the weather unit was pertinent to an upcoming unit on plants, but beyond that saw few connections, explaining that, "in high school it's more specialized." Instead, Celia emphasized the value of the units we observed for building general academic skills like taking notes and setting goals, rather than specific scientific understandings. She explained how the library research paper students were assigned as part of the weather unit would help them write better papers in future courses.

Celia made slightly more frequent and consistent relevance statements than Jackie or Kate did: In any given class period we observed her making anywhere between 0 and 5 such statements, averaging about 1.7 statements per class period across all the sessions we observed (total statements coded = 47, see Table 5). The vast majority of her relevance statements (80%) were of the "low specificity" variety, meaning that Celia rarely connected science content to outcomes, goals, or decision making, but instead pointed out the existence of scientific constructs in students' daily experience (like when she asked students if they recognized various types of clouds in the sky as they were leaving school the previous day). By way of comparison, these "low specificity" statements accounted for only10-30% of the statements made by the other teachers. The types of statements we observed in Celia's classrooms seem to embody her strategy of "making it personal."

Celia noted student ability as a primary determinant of which students might have a future in science, though she framed her discussion of ability in a different way than Jackie did. She described all of the students in her third period class as having no future in science at all, referring to them collectively as "goofballs and special education students" (we note that 6 students in this class were identified as having special education status – a number comparable to other teachers' classrooms). When probed, Celia did go on to identify two students from other class sections (one male and one female) who might have a future in science, saying they had "the inquiry"; which she explained meant they asked a lot of questions.

Celia consistently made fewer relevance statements (average of 0.6) in her 3rd period class relative to her later classes (averaging 2.3 statements in each). Nothing in Celia's interview or in our observations suggested a potential explanation for this difference in her practice, and her 3rd period class was not the one she described as being comprised of "goofballs and special education students." Thus, while Celia stated explicitly that her classroom of "goofballs and special education students" had no future in science, this did not seem to negatively affect the frequency with which she mentioned relevance. Recall however, that Celia's relevance statements almost exclusively focused on calling attention to the existence of science constructs in students' everyday lives rather than highlighting their utility for future goals or decisions.

Just as Celia was the second most frequent user of relevance statements, her students recorded the second highest ratings of daily science utility (49%, significantly higher than Jackie and Kate but significantly lower than Donna) and global science utility (4.8, though the difference from all other teachers did not reach statistical significance).

Kate. Kate was the only teacher who did not mention connecting content to students' lives when asked about effective teaching strategies. Instead, Kate's reflection on this topic focused on her efforts to convince students that they have the ability to understand “basic facts” (a phrase used repeatedly in her interview). Her strategies included verbal persuasion and scaffolding questions to convince students that they were competent:

> *I will ask them a question that they don’t think that they know the answer to, and they will say, ‘I don’t know,’ and then I will back up and I will ask them questions that they do know the answer to, and I will guide them to the answer that they think they don’t know the answer to. And I try to do that with a lot of kids, and then I’ll say ‘well see how you, how much you really do know? See how smart you really are?’ And I think that builds a lot of self-esteem and it teaches them how to use the knowledge that they already do have, ‘cause a lot of kids don’t realize that they can think beyond what they think they know.*

When probed specifically about strategies that may connect science to students’ lives, Kate mentioned occasionally using sports to illustrate science concepts, but quickly added that this particular approach was not especially effective because it appealed mostly to boys and to the smaller number of girls who she perceived to be interested in sports. While there was no evidence in her interview that Kate saw relevance as an effective instructional strategy, she clearly expressed the belief that science had multiple applications to daily life. However, her descriptions suggest that she saw these applications as longer-term outcomes of science learning rather than present-day facilitators of it. Kate’s narrative suggested a learning sequence in which students acquire a series of “basic facts” in science classrooms which will equip them to solve real-world problems “later in their life.” About her lesson on physics she said:

> *Physics is, um, kind of a [pause] part of life, it just is… [Students] don’t understand that it is, just effects your everyday living, but it does. If they understand the basics of physics they will better be able to problem-solve, just everyday situations that they will run into later in their life… they will be better able to – if they saw problems in a job, for example, if they are janitors, plumbers, mechanics, or any sort of occupation like that, will help them problem-solve simple [pause] problems.*

When asked to reflect about why the lessons we observed were important for students to know, Kate returned to this idea of future applicability. Whereas Jackie and Celia discussed the importance of their content almost exclusively in terms of simple everyday experiences, Kate focused on its importance for future careers. Kate’s elaboration of this point, however, suggested fairly low expectations regarding what her students might learn in science, and a relatively limited range of jobs in which they might apply this knowledge. For example, about her physics unit, Kate said: “if they become a janitor, plumber or mechanic it can help them remove a bolt.” Beyond these rather low-level connections, Kate struggled to articulate the importance of her units on weather and physics in terms of students’ present lives, their broader scientific understanding, or their future education.

Our classroom observations yielded a total of 38 relevance statements made by Kate – an average of 0.8 relevance statements per class period, which essentially makes her tied with Jackie for the least frequent use of relevance statements (see Table 5). Like Jackie, Kate’s most frequently observed “relevance count” for a given class period was 0 (though she was observed in one class period to make as many as four relevance statements). While infrequent, Kate’s relevance statements represented a broad range of connections in terms of our specificity to goals dimension. They articulated examples of science content in everyday existence (low specificity – 29%), described how content impacts the self or others for specific passive outcomes (moderate specificity – 46%) and made explicit connections between content and goals, decisions, plans, or actions (high specificity – 25%). This range in the specificity of relevance statements Kate made during instruction is consistent with the idea she expressed in her interview that science has many applications. The fact that she made such statements relatively infrequently during her interactions with students may reflect her belief that highlighting connections outside the classroom was not a particularly effective instructional strategy.

Similar to Celia and Jackie, Kate was inconsistent across her class sections in the extent to which she employed relevance in instruction. She tended to make more frequent relevance statements in her two morning classes (each with average of 1.6) than in the classes that occurred in the afternoon (average of 0.4 and 0.2, respectively). Our fieldnotes indicate that we regularly observed signs of Kate becoming increasingly fatigued as the day went on, frequently closing and rubbing her eyes, and verbally expressing frustration with students in her later class periods. In her interview, Kate described these classrooms differently in terms of students’ overall ability and learning needs. She described her first period class as “a much higher-achieving class overall.” Kate described her third period class as “having equal IQ” to her first period class but as being “slower learners,” explaining, “I think they’ve just been slipped through the cracks in their elementary years. They just haven’t been guided enough to be…to fill their potential.” She described both of her afternoon classes as being “very mixed” explaining that in her sixth period class there were just a few “higher-achieving kids that have had to wait for a lot of kids to catch up.” She continued, “I have kids that are truly, truly LD in there… and then I have some really slow learners in there. Truly, truly slow learners.” She described her seventh period class as “typical of seventh grade,” representing the gamut of abilities and needs. School records confirm that there were no students designated as having special needs in her morning classes, while there were six such students in each of her afternoon classes. Like Jackie, Kate appears to use relevance statements more frequently in class periods that she perceives to be higher achieving.

Also like Jackie, Kate’s students represent the lower end of the distribution on daily science utility (32%) and global science utility (4.4, significantly lower than Celia and Donna). Again, we see consistent patterns between her stated relevance beliefs, the relevance practices she employed in instruction, and students’ perceptions of utility value.

When asked about which of her students might have a future in science, Kate identified an “extremely bright” male student from her first period class, explaining that he asked “very high-level questions.” She also noted that he was a “curious young man,” noting that he pursued science interests in his free time through rock collections and moth collections. When prompted, she mentioned a female student as well, saying “again, she’s very, um, bright. She, again, reasons.” Her description of each of these students focused on both their ability and their curiosity.

Donna. Donna’s profile stands out from the others in terms of both the breadth and depth of her beliefs about and expressions of relevance. Like Jackie and Celia, Donna expressed the belief that connecting science content to students’ experience was an effective instructional strategy. She described how, in a lesson on inertia, students need to observe objects they were familiar with like bowling balls and rubber balls rolling down a ramp, explaining how demonstrating scientific principles using objects from their everyday life helped students understand abstract ideas. But Donna’s response to questions about why science content was important indicated a sophisticated understanding of why all students needed science. She began by describing her role as a science educator in terms of establishing a foundation for students’ future science education, and also expressed the desire to help her students use what they are studying for practical purposes. When they were studying speed and acceleration she wanted the students to learn “What is a speedometer really telling you?” noting her desire to use science to help students understand “how things work.” She connected the units we observed to students’ daily lives, to current events, to experiences from her own life, to natural phenomena, to students’ future education, and to potential future careers. For example, she noted that many of her students looked forward to being able to drive a car, and that Newton's laws were relevant for this activity. She talked about how the units we observed were relevant to satellite launches (related to a current news story), to sports her students were interested in, and to daily functioning like the friction that impeded a chair shoved in anger from sliding throughout the remainder of the school year. She talked about the importance of her units to understanding biology and chemistry in high school and how both chemistry and physics were related to the human body and even to the goal of curing cancer. In short, relative to the other teachers, Donna saw the content she was teaching as having much broader and complex application and meaning outside of the classroom learning context.

Donna's ability to articulate relevance was reflected in her teaching. She referred to relevance an average of 4.6 times in each class session – far more frequently than the other teachers (183 total statements, 0–7 per class period, see Table 5). While the frequency of Donna's relevance statements varied somewhat from day to day and from class period to class period, we did not detect any consistent trends across the day as we did with other teachers. Not only did Donna make more frequent relevance statements, she also made a broad range of statements in terms of specificity to goals (21% low, 41% moderate, 38% high).

Donna's students appeared to recognize her frequent usage of relevance statements as something unique and even amusing, often commenting on how she had to "bring science into everything" (to which she once replied "of course!" with an enthusiastic smile). For example, one day a student opened a can of carbonated water in class and it exploded, spewing water everywhere. As she told her student very publicly to clean up the mess, she launched into a scientific description of how the displacement of the gas (and water) from the can into the classroom was a necessary consequence given the conditions inside the can. Students in the class laughed at this response, and several rolled their eyes making the comment to one another, "there she goes again, gettin' all science-y on us." We never observed this kind of explicit awareness of teachers' use of relevance strategies by students in Jackie's, Celia's, or Kate's classes.

Unlike her peers, Donna also made statements that communicated high expectations about how students might employ their scientific knowledge in the future. When students offered correct answers to her questions about physics, for example, Donna responded with comments such as "OK. Now I know who my rocket scientists are." She once announced that she would hang a large portrait of one of her female students in the school when that student made important contributions to space travel in the future. When explaining shifts away from government funding of space travel, Donna said to her students "We don't have the science right now to do a lot of these things, so there's a lot for you guys to do." So not only did Donna regularly point out (often in very ordinary ways) how all aspects of students' life were related to science, she also expressed the expectation that her students would play an important role in scientific advancement. This latter type of statement was simply not observed in any of the other teachers' classrooms.

Our observations of Donna's interactions with students are consistent with the idea expressed in her interview that science was for everyone. In contrast to her peers, Donna said that many of her students had futures in science and her description about who these students were focused entirely on interest and curiosity. She saw these qualities as paramount in predicting a student's future in science, she said, because they will lead students to pay attention and to persist in studying. She then went on to describe her students and her content area in a way that suggested she believed that most if not all of her students could and would apply their science knowledge in a range of ways both now and in the future. She discussed how current events, sports, driving, cars, the brain and the environment were all things that were both explained by science and were naturally interesting to the students. Her belief was that all students tended to have natural curiosity about the world around them, and that her job as a science educator was to help them develop the tools to explore these curiosities in increasingly complex ways throughout their lives. When pressed to mention specific students who might have potential careers in science (her initial response to this question was "there are so many"), she mentioned 8 different students (boys and girls), and in turn described each one of them as always wanting to know "why," providing several specific examples:

> *Well, why does it happen? Why? Why… what is it that makes a snake's… a cobra's hood come out? Is that a bone in there?… Why, why? [they] always want to know. Yeah, but… well what if? What if it was like this? What if we just changed this part? Why do cells have a nucleus? Could something live without a nucleus? The 'but why' question is, you know, the way they want to go on trying to figure it out, and this makes them scientific material… when people ask those kinds of questions, if conditions are right, they go into science.*

Given the views expressed by the three other teacher participants, it is noteworthy that in describing students' future in science, Donna made no reference to students' ability, intelligence, special education status, or status as gifted.

Though Donna made no spontaneous mention of gender differences in response to other questions, when probed about it she responded that her male students were generally more interested in science than her female students "because I have more males asking me questions of "why?" and I have less females. She described how she used questioning as a measure of interest, noting that girls asked more questions during the weather unit than they did at the beginning of the physics unit until she shifted the kinds of connections she made:

> *The guys were kind of getting into the whole idea of sports. That you put a certain amount of force onto it, a certain amount of this, and the girls were less into that. But then they all were interested in cars. So that was kind of the leveling point, was the cars, 'cause they, they all want to drive cars, they all want to do that. So, it's interesting. [And next] we're getting into, we're talking about cells, you know, your brain, and how your brain is made up of cells and stuff like that, and how women have two communication centers and how it's a genetic thing, that it's really that, and the girls are, like, perking up about this, that there's this difference just because you're a girl…so they're asking questions.*

This reflection suggests that Donna's use of frequent and varied relevance statements is very much an intentional strategy aimed at capitalizing on the interest and curiosity of the broadest possible range of students.

4. Discussion

When considered as parallel sources of information, the qualitative and quantitative data strands yield results that complement one another. Our classroom observations document teachers' attempts to connect course content to their students' lives, and these observations are enhanced and triangulated by the teacher interviews, (addressing research question 1). Alone, these rich qualitative descriptions of teachers' beliefs and actions do not provide access to how teachers' efforts to communicate relevance are received by the students in their classrooms. The survey data suggest that students' beliefs about the utility value of science are variable and systematically related to who one's teacher is (addressing research question 2), but do not provide us with many clues about how or why this variation may occur. These multiple data sources are more powerful and meaningful together than alone. We employed two distinct analytic strategies to integrate our data strands in order to address our third research question. Quantizing our classroom observations and merging them with student survey data allowed us to examine general relationships between what teachers say in class and how students value the content they are learning, while statistically controlling on other factors that might influence students' value perceptions. Constructing teacher profiles helped us to understand the unique ways these relationships develop and unfold in particular teachers' classrooms. Together, these integrated approaches shed light on the complex process of how teachers shape students' appreciation of what they are learning in science.

4.1. Some teachers don't say much about relevance, but what they do say matters

The four teachers in our study differed markedly from one another in terms of how often they talked about relevance in their classrooms. Jackie and Kate arguably made infrequent attempts to connect their content to issues, competencies and concerns of the "real world" – making such connections fewer than once per class session. Donna on the other hand made such connections approximately once every 10 min across all of her classes, and Celia fell somewhere in the middle. Despite these observed differences between teachers, the merged analysis suggests that "if teachers say it, students see it." On any given day, students were more likely to see that day's course content as being useful if that day's lesson included more relevance statements. Similarly, the more a teacher maintained consistent usage of relevance statements over time, the more her students increased in perceived global science utility (though this effect was statistically significant only among girls). Importantly, the association between teachers' relevance statements and students' utility perceptions was significant even after controlling for students' perceived level of competence and interest, all of which are presumed to play influential roles in how utility information is perceived (see Durik et al., 2015; Steingut, Patall & Trimble, 2017 for reviews). Our results suggest that teachers' relevance statements shape students' utility perceptions rather broadly and are influential across students with varying levels of perceived competence and interest.

In addition to between-teacher differences in the use of relevance statements, we also noted within-teacher variation from day to day and from class period to class period within the same day. This makes sense, of course: teachers are human after all and are not wired to emit uniform "doses" of relevance each day or follow exactly the same teaching script when covering the same topic with multiple course sections. This means that from day to day and course section to course section, students may have different experiences of course-related relevance, even if they have the same teacher. What is somewhat troubling about this is that, for some teachers at least, the degree to which they emphasized relevance may be related to broad judgements teachers make about a classroom of students. This issue will be taken up in greater depth momentarily.

The fact that the teacher herself remained a significant predictor of daily utility even when her frequency of relevance statements was accounted for suggests that there were other differences in how Donna, Celia, Kate and Jackie presented content that fall outside of our definition of relevance, but that nonetheless appear to shape students' daily utility perceptions. So, while teacher relevance statements appear to heighten students' daily utility perceptions, they are not the only influence teachers have on these perceptions. It might be fruitful in future research to explore whether teacher statements about students' competence have any impact on their students' utility perceptions, as our results suggest that the relation between students' perceived competence and their perceptions of daily utility varied by classroom. The teacher profiles are also suggestive that teachers held varied views about their students' abilities, and that these views may play an important role in how teachers frame science content.

The more robust effects of both relevance statements and teacher on students' daily utility relative to global utility suggests that teachers can have relatively immediate effects on students' beliefs about specific learning tasks (i.e. "that thing we did today is useful"), and that the level of emphasis on relevance that we observed in this study was sufficient to see these effects. However, the more muted effects for both teacher and relevance statements on global utility suggest that it may be more difficult to change students' beliefs about a discipline (i.e., "science in general is useful"). This kind of change may require more time (this study lasted only seven weeks), greater frequency and consistency of relevance statements than was observable with this sample of teachers, or something else altogether.

4.2. Potential barriers (and gateways) to effectively communicating relevance

The teacher profiles enabled us to explore the unique ways that teachers' beliefs about their content, their students, and their role as educators informs their instructional practice and shapes their students' beliefs about the value of course content. In examining the interplay of these factors in multiple classrooms, one can begin to extract some ideas about the conditions that may facilitate or hinder teachers' efforts to connect content to their students' lives. In their interviews, all of the teachers except Kate spontaneously mentioned relevance (or "making it personal") as a highly effective strategy, even though we didn't ask specifically about using relevance in instruction. This raises the question then, if teachers recognize relevance as a highly effective instructional strategy, why didn't Jackie or Kate or even Celia spend more time actually trying to communicate relevance to their students? Scholars who are involved in teacher preparation have lamented that the large number and variety of prominent motivation constructs, coupled with the limited time devoted to motivation in teacher preparation programs, are to blame for future teachers who are often "unable to integrate the teaching strategies for student motivation into a meaningful and practical model that they could or would use as teachers" (Jones, 2016 p. 126, see also Wentzel & Wigfield, 2009). Our integrated analyses of multiple data sources helped us to generate a number of ideas about why some of the teachers didn't emphasize relevance all that often during instruction (and conversely, why Donna, specifically, did). While supported by our results, each of these potential explanations will require additional research to move them beyond informed speculation.

Knowledge and experience. Teachers may infrequently communicate relevance to students because they themselves do not fully understand the ways in which their content is meaningful or useful. Three out of the four teachers in the study struggled to articulate to their interviewer why it was important for students to learn course material. While it is possible that Donna invested more effort in the task of identifying connections to students' lives as a pedagogical strategy, it is more likely that Donna simply had deeper knowledge of the science content she was teaching, which enabled her to more readily make these connections. Donna had considerably more life experience and more teaching experience than Celia and Jackie and had more formal education than all three of the other teachers. Her advanced education may have helped Donna develop deeper content knowledge in the sciences as well as deeper pedagogical content knowledge, both of which may have enhanced her effectiveness as an educator in multiple ways. We also learned that Donna had read fairly extensively about fixed and incremental views of ability (or growth mindset, see Dweck, 1999) prior to our study: On the surface, Donna was the only teacher who didn't appear to have a fixed mindset regarding her students, and this seemed more likely due to her prior experience than the classroom interventions that were implemented alongside our observations. Kate arguably had educational and professional experiences in the sciences that rivaled Donna's but may have had less experience in pedagogy. This experiential difference may have made her less able to translate her content knowledge into meaningful learning experiences for her students, particularly across a broad range of content areas.

Our observations of Celia, who was one of our less experienced teachers, suggests that she may have been learning from her own experience within the span of a teaching day. Celia consistently made very few relevance statements in her first class of the day and then made more frequent relevance statements in her later classes as the day went on, and this trend was unrelated to any student or classroom characteristics we could discern. Such a pattern suggests that she might have been developing these connections as a result of multiple attempts to teach the same information. We note that we didn't observe a similar growth trajectory in Jackie, the other less experienced teacher: Her pattern across the day seemed to be related to her perceptions of student ability, an idea to which we now turn.

Beliefs about student ability and interest. Teachers may be less likely to emphasize relevance if they do not believe their students are particularly capable of doing science. In various ways, Jackie, Celia, and Kate communicated the belief that one's science ability in seventh grade might be a good predictor of who had a future in science. Further, it was apparent that in their view, certain students had the ability while others did not. This combination of beliefs about ability may have set up a chain of reasoning in which it was not logical to spend time and effort emphasizing science relevance among students who have little chance for a future in it (see Dweck, 1999). Jackie's profile exemplifies this line of thinking, and Kate's supports it to some extent as well. For these teachers, implicit beliefs about students' ability and corresponding conclusions about their potential in science may have served as a barrier to more frequent communication of relevance. In contrast, Donna appeared to see inquisitiveness and curiosity as the markers of a future in science at this age, and she further saw these characteristics as present in almost everyone. Donna was also the only teacher who explicitly communicated high expectations (in her interview and her classroom interactions) about the contributions she expected her students to make to the field of science, which implies confidence in their abilities. Both Kate and Celia briefly alluded to the importance of curiosity and interest for students' growth in science but attributed these characteristics to a small number of students. Moreover, both teachers expressed the belief that large numbers of their students (indeed entire course sections) simply lacked ability and thus would not move forward in science. It is not our intent to discount the very real challenges associated with teaching students who have special needs or who struggle academically. Our point is that Donna's beliefs about ability, interest, and curiosity were more inclusive than exclusionary, and that this belief likely led her to emphasize relevance more frequently and more consistently across course sections than any of her peers.

It is possible, though not strongly supported by evidence, that the classroom interventions that were occurring alongside our observations may have played a role in shaping teachers' thinking about how their students' ability was related to their potential in science. Donna's interview reflections and instructional behaviors suggest an underlying belief and expectation that all students can succeed – a central component of the mindset intervention that was conducted in her classes. However, Celia's classrooms were assigned to the same interventions as Donna's, and her beliefs about ability were more consistent with the other two teachers who had not received mindset training. Donna was the only one among the three participant teachers who had previously read anything about mindset, and she was fairly knowledgeable about the subject before we started working with her. Thus, it seems likely that Donna's beliefs about student ability and any influence these beliefs might have had on how she emphasized relevance may have been a reflection of her life experiences prior to the present study rather than a result of the study design. In any case, teachers' beliefs about ability and interest appear to shape how they frame course content in terms of relevance.

Perspectives on gender (and possibly race, ethnicity, and poverty). The finding that teachers' relevance statements were associated with changes in global utility for girls but not boys was not predicted but raises questions about the role of gender in teachers' communication of relevance. This finding is particularly compelling in light of the fact that the teachers in our study largely felt that their female students didn't connect to science to the same extent males did. In a recent meta-analysis, Steingut et al. (2017) report that the positive effects of rationale provision on subjective task value were larger in studies that had a greater proportion of females. The authors suggest that the gender-related effects implied by this pattern might be explained by females' greater need for affiliation, citing the need for more research to examine this hypothesis. Perhaps relative to males, the female students in this study did not personally identify as strongly with science to begin with, so were more strongly influenced when presented with information explicitly pointing out connections to their lives. There is abundant research documenting that, at all age levels, science is viewed primarily as a "male" profession (see most recently, Miller et al., 2015; Wang & Degol, 2017). Perhaps when seventh grade girls are exposed to information that contradicts this stereotype by highlighting the relevance of science to their life, their perceptions of utility change.

That the girls in this study appear to be more sensitive than boys to teachers' utility statements is interesting given that all of the teachers in the study were women. Given that models tend to be more effective if learners perceive them as similar to themselves (Bandura, 1986), male students may have been less swayed by their female teachers' comments, though this particular question would have to be explored in future studies that include male and female teachers. While all of these gender-related findings are worthy of additional investigation in the future, we want to caution the reader against drawing sweeping conclusions about the role gender played in teachers' and students' thoughts and actions. We saw no evidence in our observational data that teachers talked to male and female students differently about relevance (primarily because teachers almost always addressed relevance statements to the entire class), and most of the teacher reflections regarding gender came about as a result of specific questions asking them if they noticed gender differences. Thus, we do not have particularly compelling evidence to explain the different associations by gender in terms of a sequence of teacher beliefs and actions.

While issues of race, ethnicity and poverty were never mentioned explicitly in our interviews with any of the teachers and were not intended to be major foci of the study, several things led us to wonder about their role in how teachers thought about and enacted relevance, as well as how they might shape students' perceptions of utility. There is longstanding underrepresentation of persons of color in STEM fields, (Adamuti-Trache & Sweet, 2014; Riegle-Crumb, Moore & Ramos-Wada, 2011). While preliminary analyses showed that ethnicity was not significantly associated with students' utility perceptions in any of the quantitative and merged analyses, this may have been due to a certain degree of homogeneity in our sample. The majority of students in our sample identified as Hispanic and/or Black, and similarly high percentages of students received free or reduced-price lunch. Students from minoritized groups may identify less strongly with science goals because they have been marginalized from meaningful engagement with science at many levels (Barton et al., 2013; Bouchey & Harter, 2005), because they experience fewer similar role models in science (Roderick, 2003), or because they are less frequently encouraged to pursue science-related paths (Bouchey & Harter, 2005; Oyserman, Bybee, & Terry, 2006; Pizzolato, 2006). All of the teacher participants were non-Hispanic White, as were the vast majority of teachers in both of their schools. A theme that emerged from our teacher profiles had to do with beliefs about who science was "for" and who science was "not for." To a certain extent such beliefs may reflect white, middle class expectations about students, families, teaching, and learning. For example, teachers' implicit beliefs about ability (discussed previously) might be informed by their particular sociocultural position. Jackie's beliefs about how students' parents were involved in their education, and how this related to their potential in science reflected a white middle class model of parenting. These topics should be taken up in future research on relevance.

4.3. Limitations

This study is limited in several respects. First, it involved only 4 teachers in one school district whose student population was comprised primarily of Hispanic students from lower-income families, so our results may not generalize beyond this population (though generalization was not necessarily a goal). The diversity in experience of the four teacher participants is both a strength and a challenge for making sense of our data. It is likely that all of these differences help explain why teachers differently used relevance statements and add to the richness of how relevance information is understood and conveyed. However, this diversity created challenges for telling one story that was informed by data, as there seemed to be many.

Second, because relevance statements occurred relatively rarely, we could not go beyond a basic frequency count to quantitatively analyze links between different types of relevance statements and students' perceptions. For example, we would have liked to examine whether the level of specificity of goal connectedness in teachers' statements was associated with students' perceptions, but this was not possible.

Third, a few features of the study design have the potential to compromise the certainty of our conclusions. We took every possible step to ensure that our findings were not attributable to these design features, but we would be remiss if we did not acknowledge them here. Because the observational data involved field notes rather than video recordings, it is possible that we missed some teacher relevance statements and have underrepresented this teacher practice. We attempted to mitigate this threat to credibility by having multiple observers present in every classroom we observed, but it is possible that certain statements could have slipped by our notice, particularly on the relatively infrequent occasions where we had two observers rather than three. Due to the large number of class sessions observed it was not possible to have the same team of three observers present in every single classroom. We attempted to mitigate this threat to the trustworthiness of our fieldnotes by ensuring that at least one of the study's Principal Investigators (first and fourth authors of this article) were present at all times. Additionally, the classrooms in the present study were participating in a number of classroom interventions alongside our observations, which complicated our analysis process considerably. In all phases we took care to examine whether there was any evidence that this study design element played a role in the statements teachers made, in students' utility perceptions, or in the relation between the two. In the quantitative and merged analyses, we included intervention group in all of the models and it never emerged as a significant predictor of students' utility perceptions, frequency of teacher's relevance statements, or as a moderator of the relation between the two. It is also possible, though not strongly supported by available evidence, that the study design may have played a role in shaping teachers' thinking about which students had potential in science. Nonetheless, the fact that teachers had been assigned to these intervention conditions is a potential limitation and future studies will ideally not have this added complication.

Fourth, while the design of the study provided a relatively fine-grained description of the degree to which a given teacher emphasized relevance from day to day from one class period to another, we are somewhat limited in our ability to confidently draw many conclusions at this time about why this variation exists. Our interviews and observations were suggestive of some broad reasons a teacher may or may not use relevance (e.g., Jackie may only emphasize relevance in her classes that have large numbers of gifted students, while Kate tends to avoid relevance statements in classes with many special education students), but the reasons for much of the day-to-day variation we observed are still unknown at this time.

Finally, the study's design afforded few opportunities to consider how issues of race, poverty, and their intersection with gender may have influenced teachers' relevance statements and students' interpretation of them. These design constraints may limit the interpretation of the data. It is particularly important to acknowledge that like most theories in educational psychology, the theoretical models regarding relevance and task value that frame this study were developed without consideration of race as a socio-historical construct, and as such afford an approach to race that is limited to explaining variance and comparing groups (DeCuir-Gunby & Schutz, 2014). DeCuir-Gunby & Schutz (2014) propose that scholarship examining race advance *race-reimaged constructs*, which involves making race a central focus of the study rather than a source of variation. A race-reimaged approach involves viewing traditional constructs like task value and relevance from a socioculturally relevant lens, like that which underlies Critical Race Theory (Delgado & Stefancic, 2017). In this approach the focus is on how racial and ethnic groups experience particular constructs. While the current study is not positioned to do such re-imaging due to the study's focus and the nature of the data collected, it will be important to bring this important critical perspective to future studies of relevance, particularly when there is a mismatch between students' and teachers' sociocultural groups.

4.4. Conclusions

Brophy (2008) laments the line of thinking, common among teachers and parents alike, that even though students may not appreciate the value of a learning experience now, they will see its value in the future. He calls this thinking misguided, saying that more effort should be focused on how teachers can help students see the value of what they are learning while they are learning it. The results of this mixed methods study suggest that teacher relevance statements may be a step in the right direction toward this goal. When science teachers verbally emphasize the relevance of their course content during their everyday instruction, their students are more likely to see the utility of their immediate science tasks. Moreover, when teachers make such relevance statements consistently over time, female students in particular develop enhanced perceptions of the utility value of science as a domain. Teachers' ability and willingness to integrate relevance into their instruction may depend on how deeply they understand the content, and on the beliefs and expectations they hold regarding their students. While utility value represents but one of many ways students can come to value content within a domain, this study adds to our understanding of how teachers can support and nurture students' appreciation for what they are learning while they are learning it. Given the links between utility value perceptions and other academic and motivational outcomes (Eccles et al., 1983, 1998; Lau & Roeser, 2002; Hulleman & Harackiewicz, 2009; Hulleman et al., 2008; Simpkins et al., 2006), encouraging teachers to make more frequent and consistent references to relevance as a matter of course in their instruction could have substantial payoffs in both the short and long term.

Acknowledgements

This material is based upon work supported by the National Science Foundation, United States under Grant No: HRD-1136143. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the National Science Foundation, United States. We thank the teachers, students, and schools who participated in this research, and the graduate research assistants who helped us collect and code this large corpus of data. Finally, we are grateful to the editors of this special issue and the anonymous reviewers for their helpful comments on previous drafts of this article.

Appendix A. Description of intervention groups that ran concurrently to the classroom observations

A1. Summary writing condition

Four course sections taught by two teachers were assigned to a summary writing condition. Once a week for six weeks, students wrote brief (5-sentence) essays summarizing what they were learning in science that week (writing task developed by Hulleman & Harackiewicz, 2009). Summary essays were collected by the researchers: Teachers were not in the room when initial instructions were given, were unaware of what students were being asked to write about and did not see the essays.

A2. Relevance writing condition

Three course sections taught by two teachers were assigned to the relevance writing condition. Once a week for six weeks, students wrote brief (5-sentence) essays about how what they were learning in class that week related to their lives or the lives of someone they know (developed by Hulleman & Harackiewicz, 2009). Relevance essays were collected by the researchers: Teachers were not in the room when initial instructions were given, were unaware of what students were being asked to write about and did not see the essays.

A3. Mindset condition

Two teachers had two course sections each randomly assigned to a mindset only condition, for a total of four course sections. Once per week for six weeks, students participated in the Brainology® intervention – an interactive online software program designed to teach students about the science of the brain in order to support the belief that ability is incremental and malleable (growth) as opposed to fixed (MindsetWorks, 2012; https://www.mindsetworks.com). The six-week Brainology intervention was led by researchers, though teachers were present during its administration. As part of this intervention, participating teachers engaged in approximately two hours of professional development in which researchers described fixed and incremental views of ability, and the Brainology program.

A4. Relevance writing + mindset condition

The teachers whose students participated in the mindset condition had additional course sections in which students simultaneously participated in both the mindset and relevance writing interventions as described above. Three classrooms total were assigned to this condition.

Assignment of condition by teacher is presented in Table A1.

Table A1
Assignment of treatment groups across middle schools and teachers.

	School 1		School 2		
Condition	Celia	Jackie	Donna	Kate	Total
Summary writing		2		2	4
Relevance writing		1		2	3
Mindset only	2		2		4
Mindset + relevance writing	1		2		3
Total	3	3	4	4	14

Note. Numbers in each cell indicate the number of classrooms assigned to this treatment condition.

A5. A note about analysis

In all of the quantitative and merged analyses, we tested models with intervention condition entered at the level of the classroom, and intervention condition was not a significant predictor of students' daily utility or changes in students' global utility. We tested for effects of the four study conditions against each other in addition to testing for any relevance writing vs. no relevance writing (i.e, the relevance writing and the mindset + relevance writing conditions were compared to the summary writing and mindset only conditions). While this null finding may be unexpected in the context of the intervention component of the study it enables us to examine potential effects of teachers' relevance statements across study conditions because the intervention had no measurable effect on utility perceptions. Results regarding whether relevance writing interventions actually impact students' utility perceptions are somewhat mixed, and some researchers have suggested that such interventions may impact educational outcomes through mechanisms other than utility perceptions *per se*, or that a variety of factors moderate the effects of such interventions (see Steingut et al., 2017 for a recent review). A more thorough discussion the null effects of intervention condition on perceived utility in the present study is beyond the scope of this article, but is presented and discussed in Durik et al. (2018). What is important to the present study is that intervention group was not associated with either teachers' relevance statements or students' utility perceptions in any of the analyses, but teacher was. Thus, teacher, rather than intervention condition was included in final analyses.

Appendix B. Research practices employed to ensure trustworthiness in qualitative data

Guba's (1981) Criteria	Shenton's (2004) Strategies	Research practices employed in the current study
Credibility		
	Early familiarity with organizational culture	PIs visited, observed, & met w/teachers multiple times prior to study and had prolonged engagement at the research sites
	Triangulation	Interview and observational data collected to assess teachers' enactment of relevance. Multiple researchers with differing backgrounds were involved in observation and coding, minimizing potential for systematic bias
	Facilitate honesty	Established rapport with teachers through reflective listening and honoring their requests. Recognized them as "experts" on their curriculum and students
	Debriefing sessions	Researchers debriefed once per week or more
	Negative case Examination	Reread interviews after highlighting and coding to reconfirm nothing was missed
	Peer scrutiny	Advisory board of established scholars and practitioners reviewed coding frames & conclusions
	Frame findings w/prior research	Extensive review and consideration of relevant prior research undertaken
	Member checks	On the spot reflection to participant by interviewer. (we note that this represents very minimal member checking)
	Thick description	Use of quotations by participants
	Background, Qualifications of investigators. Funding agency identified	PIs had extensive experience in interviewing and classroom observation. The National Science Foundation is acknowledged as the funder
Transferability		
	Description of context, methods, & participants	Detailed descriptions are provided as space allows in this manuscript. Extensive documentation is available upon request
Dependability		
	Overlapping methods	Field notes and interviews documented classroom practices
	Describe methods in depth	Provided as space allows. All protocols, manuals, and instruments available upon request
Confirmability		
	Triangulation	See above
	Recognition of shortcomings	Limitations are included in the discussion section
	Audit trail	See Fig. 1 describing data collection sequence

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cedpsych.2018.08.005.

References

Adamuti-Trache, M., & Sweet, R. (2014). Science, technology, engineering and math readiness: Ethno-linguistic and gender differences in high-school course selection patterns. *International Journal of Science Education, 36*, 610–634. https://doi.org/10.1080/09500693.2013.819453.

Anderman, L. H., Andrzejewski, C. E., & Allen, J. (2011). How do teachers support students' motivation and learning in their classrooms. *Teachers College Record, 113*, 969–1003.

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology, 72*, 261–278. https://doi.org/10.1348/000709902158883.

Bandura A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall. Doi: 10.5465/amr.1987.4306538.

Barton, A. C., Kang, H., Tan, E., O'Neill, T. B., Bautista-Guerra, J., & Brecklin, C. (2013). Crafting a future in science: Tracking middle school girls' identity work over time and space. *American Educational Research Journal, 50*, 37–75. https://doi.org/10.3102/0002831212458142.

Bergin, D. A. (1999). Influences on classroom interest. *Educational Psychologist, 34*, 87–98. https://doi.org/10.1207/s15326985ep3402_2.

Bouchey, H. A., & Harter, S. (2005). Reflected appraisals, academic self-perceptions, and math/science performance during early adolescence. *Journal of Educational Psychology, 97*, 673–686. https://doi.org/10.1037/0022-0663.97.4.673.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101. https://doi.org/10.1191/1478088706qp063oa.

Brophy, J. (2008). Developing students' appreciation for what is taught in school. *Educational Psychologist, 43*, 132–141. https://doi.org/10.1080/00461520701756511.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*, 97–114. https://doi.org/10.1177/1468794106058877.

Clegg T., Bonsignore E., Ahn J., Yip J., Pauw D., Gubbels M., ... Rhodes E. (2014). Capturing personal and social science: Technology for integrating the building blocks of disposition. In Proceedings of the eleventh international conference of the learning sciences boulder, Colorado, USA, pp. 455–462.

Clegg, T., & Kolodner, J. (2014). Scientizing and cooking: Helping middle-school learners develop scientific dispositions. *Science Education, 98*, 36–63. https://doi.org/10.1002/sce.21083.

Creswell, J. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J., & Plano Clark, V. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Disease, 175*, 526–536. https://doi.org/10.1007/978-94-017-9088-8_3.

DeCuir-Gunby, J. T., & Schutz, P. A. (2014). Researching race within educational psychology contexts. *Educational Psychologist, 49*, 244–260. https://doi.org/10.1080/00461520.2014.957828.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*, 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x.

Delgado, R., & Stefancic, J. (2017). *Critical race theory: An introduction* (3rd ed.). NYU Press.

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology, 99*, 597–610. https://doi.org/10.1037/0022-0663.99.3.597.

Durik, A. M., Hulleman, C. S., & Harackiewicz, J. M. (2015). One size fits some: Instructional enhancements to promote interest. In K. A. Renninger, & M. Nieswandt (Eds.). *Interest, in mathematics and science learning* (pp. 49–62). Washington, DC: American Educational Research Association.

Durik, A. M., Schwartz, J., Schmidt, J. A., & Shumow, L. (2018). Age differences in effects of self-generated utility among black and hispanic adolescents. *Journal of Applied Developmental Psychology, 54*, 60–68. https://doi.org/10.1016/j.appdev.2017.11.004.

Durik, A. M., Shechter, O. G., Noh, M., Rozek, C. S., & Harackiewicz, J. M. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interestand performance. *Motivation and Emotion, 39*, 104–118. https://doi.org/10.1007/s11031-014-9419-0.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Taylor & Francis.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values, and academics behaviors. In J. T. Spence (Ed.). *Achievement and achievement motives* (pp. 75–146). San Francisco, CA: Freeman.

Eccles, J. S., Barber, B. L., Updegraff, K., & O'Brien, K. M. (1998). An expectancy-value model of achievement choices: The role of ability self-concepts, perceived task utility, and interest in predicting activity choice and course enrollment. In L. Hoffman, A. Krapp, K. A. Renninger, & J. Baumert (Eds.). *Interest and learning: Proceedings of the Seeon conference on interest and gender* (pp. 267–280). Keil, Germany: IPN.

Frymier, A. B., & Shulman, G. M. (1995). "What's in it for me?": Increasing content relevance to enhance students' motivation. *Communication Education, 44*, 40–50. https://doi.org/10.1080/03634529509378996.

George, R. (2000). Measuring change in students' attitudes toward science over time: An application of latent variable growth modeling. Retrieved from *Journal of Science Education and Technology, 9*, 213–225. https://link.springer.com/content/pdf/10.1023/A:1009491500456.pdf.

Giorgi, A. (1985). *Phenomenology and psychological research.* Pittsburgh, PA: Duquesne.

Giorgi, A. (1994). A phenomenological perspective on certain qualitative research methods. *Journal of Phenomenological Psychology, 25*, 190–220. https://doi.org/10.1163/156916294x00034.

Gnambs, T., & Hanfstingl, B. (2016). The decline of academic motivation during adolescence: An accelerated longitudinal cohort analysis on the effect of psychological need satisfaction. *Educational Psychology, 36*, 1691–1705. https://doi.org/10.1080/01443410.2015.1113236.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*, 3–13. https://doi.org/10.1037//0022-0663.93.1.3.

Greenfield, T. A. (1997). Gender- and grade-level differences in science interest and participation. *Science Education, 81*, 259–276. https://doi.org/10.1002/(sici)1098-237x(199706)81:3 < 259::aid-sce1 > 3.3.co;2-3.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. Retrieved from *Educational Technology Research and Development, 29*, 75–91. https://link.springer.com/content/pdf/10.1007/BF02766777.pdf.

Greene, J. C. (2007). *Mixed methods in social inquiry.* San Francisco: Jossey-Bass.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*, 255–274. https://doi.org/10.2307/1163620.

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, E. A., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest and performance. *Journal of Educational Psychology, 100*, 105–122. https://doi.org/10.1037/0022-0663.100.1.105.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life.* Thousand Oaks, CA: Sage.

Hidi, S., & Renninger, A. K. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127. https://doi.org/10.1207/s15326985ep4102_4.

Hill, C., Corbett, C., & St Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* Washington, DC: American Association of University Women.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*, 1277–1288. https://doi.org/10.1177/1049732305276687.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*, 1410–1412. https://doi.org/10.1126/science.1177067.

Hulleman, C. S., Durik, A. M., Schweigert, S. B., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*, 398–416. https://doi.org/10.1037/0022-0663.100.2.398.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*, 880–895. https://doi.org/10.1037/a0019506l.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509–527. https://doi.org/10.1111/1467-8624.00421.

Jang, H. (2008). Supporting students' motivation, engagement, and learning during an uninteresting activity. *Journal of Educational Psychology, 100*, 798–811. https://doi.org/10.1037/a0012841.

Jones, B. D. (2016). Teaching motivation strategies using the MUSIC model of motivation as a conceptual framework. In M. C. Smith, & N. Defrates-Dench (Eds.). *Challenges and innovations in educational psychology teachinga and learning* (pp. 123–136). Charlotte, NC: Information Age.

Kafkas, S. S. (2016). *Why do we need to learn this? An investigation of the association between science teachers' use of utility value statements and students' immediate and global perceptions of science utility. Unpublished doctoral dissertation.* DeKalb, IL: Northern Illinois University.

Keller, J. M. (1987a). Development and use of the ARCS model of instructional design. *Journal of Instructional Development, 10*, 2–10. https://doi.org/10.1007/bf02905780.

Keller, J. M. (1987b). Motivational design. In R. McAleese, & D. Unwin (Eds.). *Encyclopedia of educational media, communications, and technology*(2nd ed.). Westport, CT: Greenwood Press.

Keller, M. M., Woolfolk Hoy, A., Goetz, T., & Frenzel, A. C. (2016). Teacher enthusiasm: Reviewing and redefining a complex construct. *Educational Psychology Review, 28*, 743–769. https://doi.org/10.1007/s10648-015-9354-y.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life.* Berkeley, CA: University of California Press.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*, 139–162. https://doi.org/10.1207/s15326977ea0802_04.

Marchand, G. C., & Gutierrez, A. P. (2017). Processes involving perceived instructional support, task value, and engagement in graduate education. *Journal of Experimental Education, 85*, 87–106. https://doi.org/10.1080/00220973.2015.1107522.

Means, T. B., Jonassen, D. H., & Dwyer, F. M. (1997). Enhancing relevance: Embedded ARCS strategies vs. purpose. *Educational Technology Research and Development, 45*, 5–17. https://doi.org/10.1007/bf02299610.

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology, 107*, 631–644. https://doi.org/10.1037/edu0000005.

MindsetWorks (2012). Evidence of impact: Brainology and MindsetWorks SchoolKit. Retrieved from https://www.mindsetworks.com/websitemedia/info/impactsummaryhandout.pdf.

Moustakas, C. (1994). *Phenomenological research methods.* Thousand Oaks, CA: Sage.

Muddiman, & Frymier (2009). What is relevant? Student perceptions of relevance strategies in college classrooms. *Communication Studies, 60*, 130–146. https://doi.org/10.1080/10510970902834866.

NGSS Lead States (2013). Next generation science standards: For states, by states. Washington, DC: National Academies Press.

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press.

Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. M. Newmann (Ed.). *Student engagement and achievement in American secondary schools* (pp. 11–39). New York: Teachers College Press.

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology, 8*, 375–387.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*, 1049–1079. https://doi.org/10.1080/0950069032000032199.

Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology, 91*, 188. https://doi.org/10.1037/0022-3514.91.1.188.

Patrick, B. C., Hisley, J., & Kempler, T. (2000). "What's everybody so excited about?": The effects of teacher enthusiasm on student intrinsic motivation and vitality. *Journal of Experimental Education, 68*, 217–236. https://doi.org/10.1080/00220970009600093.

Pizzolato, J. E. (2006). Achieving college student possible selves: Navigating the space between commitment and achievement of long-term identity goals. *Cultural Diversity and Ethnic Minority Psychology, 12*, 57–69. https://doi.org/10.1037/1099-9809.12.1.57.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Ravitch, S. M., & Carl, N. M. (2015). *Qualitative research: Bridging the conceptual, theoretical, and methodological.* Thousand Oaks, CA: Sage.

Reeve, J., Jang, H., Hardre, P., & Omura, M. (2002). Providing a rationale in an autonomy-supportive way as a strategy to motivate others during an uninteresting activity. Retrieved from *Motivation and Emotion, 26*, 183–207. https://link.springer.com/content/pdf/10.1023/A:1021711629417.pdf.

Reis, H. T., & Judd, C. M. (Eds.). (2000). *Handbook of research methods in social and personality psychology*. New York, NY: Cambridge University Press.

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*, 458–476. https://doi.org/10.1002/sce.20431.

Roderick, M. (2003). What's happening to the boys? Early high school experiences and school outcomes among African American male adolescents in Chicago. *Urban Education, 38*, 538–607. https://doi.org/10.1177/0042085903256221.

Rothwell J. (2013). The hidden STEM economy. Washington DC: Brookings Institution. Retrieved 7/17/2017 from https://www.brookings.edu/research/the-hidden-stem-economy/.

Schreiner, C., & Sjøberg, S. (2004). Sowing the seeds of ROSE. Background, rationale, questionnaire development and data collection for ROSE (the relevance of science education)–a comparative study of students' views of science and science education. Acta Didactica, 4/2004. Retrieved 7/17/2017 from http://www.ils.uio.no/forskning/publikasjoner/actadidactica/index.html.

Schweinle, A., Meyer, D. K., & Turner, J. C. (2006). Striking the right balance: Students' motivation and affect in elementary mathematics. *The Journal of Educational Research, 99*, 271–294. https://doi.org/10.3200/joer.99.5.271-294.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). New York, NY: Teachers College Press.

Shechter, O. G., Durik, A. M., Miyamoto, Y., & Harackiewicz, J. M. (2011). The role of utility value in achievement behavior: The importance of culture. *Personality and Social Psychology Bulletin, 37*, 303–317. https://doi.org/10.1177/0146167210396380.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*, 63–75. https://doi.org/10.3233/efi-2004-22201.

Shumow, L., & Schmidt, J. A. (2014). *Enhancing adolescents' motivation for science: Research-based strategies for teaching male and female students.* Thousand Oaks, CA: Corwin Press.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*, 70–83. https://doi.org/10.1037/0012-1649.42.1.70.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*, 571. https://doi.org/10.1037//0022-0663.85.4.571.

Steingut, R. R., Patall, E. A., & Trimble, S. S. (2017). The effect of rationale provision on motivation and performance outcomes: A meta-analysis. *Motivation Science, 3*, 19–50. https://doi.org/10.1037/mot0000039.

Thomas, A. E. (2017). Gender differences in students' physical science motivation: Are teachers' implicit cognitions another piece of the puzzle? *American Educational Research Journal, 54*, 35–58. https://doi.org/10.3102/0002831216682223.

Van Aalderen-Smeets, S. I., & van der Molen, J. H. W. (2015). Improving primary teachers' attitudes toward science by attitude-focused professional development. *Journal of Research in Science Teaching, 52*, 710–734. https://doi.org/10.1002/tea.21218.

Vilorio, D. (2014, Spring). STEM 101: Intro to tomorrow's jobs, Occupational outlook quarterly. Washington DC: Bureau of Labor Statistics.

Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review, 29*, 119–140. https://doi.org/10.1007/s10648-015-9355-x.

Wentzel, K. R., & Wigfield, A. (Eds.). (2009). *Handbook of motivation at school*New York, NY: Taylor & Francis. https://doi.org/10.4324/9781315773384.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265–310. https://doi.org/10.1016/0273-2297(92)90011-p.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. https://doi.org/10.1006/ceps.1999.1015.