# How conveying relevance boosts students' motivation to learn mathematics: Effects of teaching strategies, classmates, and scientific interventions

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

**Brigitte Maria Brisson, M.A.**
aus Augsburg

Tübingen
2018

| | |
|---|---|
| Tag der mündlichen Prüfung: | 27.06.2018 |
| Dekan der Fakultät: | Prof. Dr. Josef Schmid |
| 1. Gutachter: | Prof. Dr. Benjamin Nagengast |
| 2. Gutachterin: | Prof. Dr. Jessika Golle |

# Acknowledgements


I would like to sincerely thank you for your invaluable support to accomplish this dissertation:

Prof. Dr. Ulrich Trautwein, Prof. Dr. Benjamin Nagengast
*as supervisors*

Prof. Dr. Chris Hulleman
*as mentor during my research stay at the University of Virginia*

Prof. Dr. Jessika Golle, Prof. Dr. Taiga Brahm
*as members of the examination committee*

Dr. Hanna Gaspard, Dr. Isabelle Häfner, Dr. Barbara Flunger, Dr. Anna-Lena Dicke
*as project colleagues, advise-givers, and friends*

Dr. Katharina Allgaier, Dr. Evelin Herbein, Cora Parrisius, Heide Kneißler, Dr. Alena Friedrich, Dr. Jessica Kornmann, Dr. Sandra Rothenbusch, Dr. Maike Sigel, Prof. Dr. Ingo Zettler, Hanh Nguyen-Zettler, Margit Fritz, Dr. Richard Göllner, Dr. Norman Rose, Dr. Wolfgang Wagner, other HIB coworkers;
Prof. Dr. Christine Bescherer, Dr. Sara Dallinger, Dr. Jürgen Schneider, Dr. Marcus Syring, other members of the cooperative research training group Uni Tübingen–PH Ludwigsburg;
Prof. Dr. Eckhard Klieme, Désirée Theis, Dr. Markus Sauerwein, other DIPF coworkers
*as colleagues, backers, friends*

1978 students and their 73 mathematics teachers
*as participants*

Dr. Gero Bauer, Dr. Hélène Caillaud, Dr. Tanja Granzow, Stefanie Kicherer, Johannes Steymans
*as flatmates and friends*

Elfriede & Bernhard Schreier
*as loving parents*

Veronika Schreier, Johanna Gottlob, Christiane Schreier, Michael Schreier, Simon Schreier
*as sisters and brothers*

Dr. Erwan Brisson
*as understanding, supportive, patient, and loving husband*

Lora Brisson
*as a special effect*

# Abstract

Mathematical skills are paramount for active participation in today's scientifically and technologically advanced society, where the number of jobs requiring mathematical and scientific literacy is growing. Although educational practitioners and politicians attribute great importance to learning mathematics at school, many students have difficulty seeing personal relevance of learning mathematics: The value or usefulness they attribute to mathematical skills decreases as does their motivation to learn mathematics as they progress through secondary school. Inability to see the importance of learning mathematics often corresponds with reduced effort, poorer self-concept and self-efficacy, and low achievement in the subject.

Helping students discover the personal relevance of developing mathematical skills is a central challenge in education. However, empirical research on the impact of everyday teaching strategies aimed at conveying the relevance of mathematics topics covered in class on students' motivation is scarce. Moreover, the role of students' perception of their classmates' mathematics-related value beliefs in the development of their own value beliefs has been neglected. While a growing number of researchers in the United States have investigated the effectiveness of scientific interventions for the classroom aimed at conveying the relevance of learning various topics, numerous questions remain unanswered. Next to the fact that relevance interventions still need to be tested outside the United States and for the subject of mathematics, study designs are needed that are highly compatible with students' genuine learning contexts. Moreover, the effectiveness of various relevance intervention approaches as measured through analysis of short-term and long-term achievement needs to be compared, and the processes underlying these effects need to be studied.

In this dissertation investigation is made into whether conveying the relevance of mathematics through (a) common instructional practices or (b) scientific intervention helps students find personal relevance of mathematical skills and thereby boosts their motivation and achievement in mathematics. To this end, three empirical studies are conducted of a sample of 1961 ninth-grade students and their 73 mathematics teachers in 25 academic track secondary schools in Germany. In Study 1 the relationship between three relevance-oriented teaching strategies and students' perception of their classmates' mathematics-related value beliefs and the students' own mathematics-related value beliefs (intrinsic, attainment, utility, and cost) and changes in those beliefs was assessed over six months. The teacher-reported strategy "introducing new mathematics topics with everyday examples" correlated with a decrease over six months in students' perceived cost of learning mathematics, and the teacher-reported strategy "demonstrating links between mathematics and other academic subjects" was associated with students' attainment value at individual assessment points but did not seem to affect the development of their attainment value over six months. Students' reports on their teachers' "stressing the practical applicability of mathematics" correlated positively with students' mathematics-related intrinsic, attainment, and utility values and negatively with their perception of the cost of learning mathematics at individual assessment points and over six months. In addition, students' perceptions of their classmates' mathematics-related value beliefs correlated positively with their own intrinsic, attainment, and utility value beliefs and negatively with cost at individual assessment points and over six months.

In Study 2 the effects of two 90-minute researcher-led relevance interventions implemented in the classroom on students' short-term and long-term competence beliefs, effort, and achievement in mathematics were investigated. Within a cluster-randomized trial, students in the experimental classes first watched a presentation on the relevance of mathematics for everyday life and career pathways and then either wrote a text about the personal relevance of mathematics ("text condition", adapted from studies conducted in the United States) or commented on statements made by slightly older peers about the relevance of mathematics for them ("quotations condition", newly developed task). Results of Study 2 show that the text condition fostered students' self-efficacy in mathematics after five months, whereas the quotations condition enhanced students' self-concept, self-efficacy, effort, and test scores in mathematics until up to five months after the intervention.

In Study 3 the possible mechanisms underlying the differences in the effectiveness of the two intervention conditions were explored by examining the antecedents and effects of students' responsiveness to the writing tasks about the relevance of learning mathematics. In both intervention conditions, highly conscientious students had comparatively high scores on the responsiveness index which assessed the degree of positive argumentation, personal connections, and in-depth reflections in students' essays. Furthermore, female students and students who initially claimed that mathematical skills were very useful responded particularly well to the text condition, whereas high achievers and students who initially were very interested in mathematics responded particularly well to the quotations condition. Comparing intervention effects on students' utility value beliefs according to students' responsiveness, highly responsive students in the text condition found mathematics more useful after the intervention, but the least responsive students found mathematics actually less useful afterwards. In contrast, in the quotations condition, both very responsive and not very responsive students profited from the relevance intervention.

The results of the three empirical studies are summarized and discussed in relation to the current state of research on students' motivation in mathematics. Implications for future research as well as educational policy and practice are deliberated.

# Zusammenfassung

Mathematikkenntnisse sind eine Voraussetzung für aktive gesellschaftliche Teilhabe in der heutigen Welt, in der technologische Neuerungen allgegenwärtig sind und die Zahl der Berufe steigt, die eine fundierte naturwissenschaftliche Grundbildung erfordern. Dementsprechend schreiben Lehrkräfte und Bildungspolitikerinnen und -politiker dem Mathematiklernen in der Schule eine hohe Bedeutung zu. Schülerinnen und Schülern hingegen fällt es schwer, Mathematikkenntnisse als persönlich relevant zu empfinden: Ihre Motivation in Mathematik nimmt im Verlauf der Sekundarschulzeit ab, insbesondere ihre Nützlichkeitsüberzeugungen in diesem Fach. Geringe Nützlichkeitsüberzeugungen gehen wiederum mit geringer Anstrengungsbereitschaft, geringen Kompetenzüberzeugungen und schwacher Leistung in Mathematik einher.

Die Schülerinnen und Schüler dabei zu unterstützen Gründe für die persönliche Relevanz von Mathematikkenntnissen zu finden ist daher eine zentrale Herausforderung der Bildungsforschung und -praxis und Hauptanliegen der vorliegenden Dissertation. Es mangelt an empirischer Forschung zur motivationalen Wirksamkeit von Unterrichtsstrategien, durch die die Relevanz von Mathematikkenntnissen vermittelt werden kann. Auch der Einfluss der Klassenkameradinnen und -kameraden und ihrer Wertschätzung für das Fach Mathematik auf die Motivation von Schülerinnen und Schülern wurde bislang kaum erforscht. Die Wirksamkeit von wissenschaftlichen Interventionen zur Vermittlung der Relevanz naturwissenschaftlicher Lerninhalte im Klassenzimmer wird hingegen zwar zunehmend erforscht, allerdings bleiben noch etliche Fragen offen. Neben der Tatsache, dass Relevanzinterventionen bisher kaum außerhalb des US-amerikanischen Kulturraums und im Fach Mathematik getestet wurden, fehlt es an Interventionsstudien mit einem Design, das dem natürlichen Lernumfeld der Schülerinnen und Schüler entspricht. Die kurz- und längerfristige Wirksamkeit verschiedener Interventionsansätze wurde bisher noch nicht umfassend untersucht, und auch die den Effekten zugrunde liegenden Wirkprozesse sind noch nicht ausreichend ergründet.

Um diese Forschungslücken zu schließen untersucht die vorliegende Dissertation, inwiefern die Relevanz der Mathematik im Klassenzimmer vermittelt werden kann durch (a) relevanzbezogene Unterrichtsmerkmale und (b) wissenschaftliche Interventionen im Mathematikunterricht. Anhand einer Stichprobe von 1978 Schülerinnen und Schülern der 9. Jahrgangsstufe und deren 73 Mathematiklehrkräfte an 25 baden-württembergischen Gymnasien wurden drei empirische Studien durchgeführt.

In Studie 1 wurde der relative Zusammenhang zwischen relevanzorientierten Merkmalen des Mathematikunterrichts (verschiedene Unterrichtsstrategien, Wertschätzung des Fachs im Klassenverband) und den Wertüberzeugungen der Schülerinnen und Schüler in Mathematik (intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugungen) erforscht. Die Unterrichtsstrategie „Themeneinführung mit Alltagsbeispielen" aus Lehrerperspektive führte zu einer Abnahme der Kostenüberzeugungen in Mathematik. Die von Lehrkräften berichtete

„Demonstration von Sachverbindungen“ im Mathematikunterricht hing positiv mit der Wichtigkeitsüberzeugung der Schülerinnen und Schüler zusammen. Die Strategie „Praxisorientierung im Mathematikunterricht“ aus Schülersicht förderte den intrinsischen Wert sowie Wichtigkeits- und Nützlichkeitsüberzeugungen und sagte eine Abnahme der Kostenüberzeugungen der Schülerinnen und Schüler über sechs Monate vorher. Die aus Schülersicht berichtete Wertschätzung der Mathematik im Klassenverband war mit einer Zunahme der Nützlichkeitsüberzeugungen der Schülerinnen und Schüler in Mathematik assoziiert.

In Studie 2 wurden die Effekte von zwei 90-minütigen Relevanzinterventionen im Klassenzimmer auf die kurz- und längerfristigen Kompetenzüberzeugungen, die Anstrengungsbereitschaft und die Leistung der Schülerinnen und Schüler im Fach Mathematik untersucht. In einem auf Klassenebene randomisierten Experiment folgten Schülerinnen und Schüler in den Interventionsklassen während einer Doppelstunde zunächst einer Präsentation zur Relevanz der Mathematik und bearbeiteten im Anschluss individuelle Schreibaufträge. Dabei kommentierten die Teilnehmenden entweder Zitate von jungen Erwachsenen zur Relevanz der Mathematik (Zitatebedingung) oder verfassten einen freien Text über die persönliche Relevanz der Mathematik (Textbedingung; adaptiert von Studien aus dem US-amerikanischen Raum). Die Zitatebedingung förderte bis zu fünf Monate lang das akademische Selbstkonzept, die Selbstwirksamkeit, die Anstrengungsbereitschaft und die Leistung der Schülerinnen und Schüler in Mathematik. Die Textbedingung förderte lediglich die Selbstwirksamkeit der Schülerinnen und Schüler in Mathematik fünf Monate nach der Intervention.

Um die Mechanismen zu erforschen, die den Unterschieden in der Wirksamkeit der beiden Interventionsansätze zugrunde liegen könnten, wurde in Studie 3 untersucht, mit welcher Qualität die Schülerinnen und Schüler die Schreibaufgaben zur Relevanz der Mathematik bearbeitet hatten. Prädiktoren und Wirkungen der Bearbeitungsqualität wurden ermittelt. In beiden Interventionsbedingungen hatten sehr gewissenhafte Schülerinnen und Schüler vergleichsweise hohe Werte auf einem Qualitätsindex, der den Grad der positiven Argumentation (d.h. für den Nutzen der Mathematik), des persönlichen Bezugs und der Reflexionstiefe in den Schüleraufsätzen maß. Leistungsstarke Schülerinnen und Schüler und an Mathematik eingangs hochinteressierte Schülerinnen und Schüler erledigten die Schreibaufgabe in der Zitatebedingung besonders hochwertig, wohingegen in der Textbedingung Mädchen und Schülerinnen und Schüler mit eingangs hohen Nützlichkeitsüberzeugungen von Mathematik den Schreibauftrag mit besonders hoher Qualität bearbeiteten. Beim Vergleich der Interventionseffekte auf die Nützlichkeitsüberzeugungen in Abhängigkeit von der Qualität der Aufgabenbearbeitung zeigte sich, dass in der Zitatebedingung sowohl Schülerinnen und Schüler mit hohen als auch solche mit niedrigen Werten auf dem Qualitätsindex von der Teilnahme an der Intervention profitierten. In der Textbedingung fanden Schülerinnen und Schüler mit hohen Werten auf dem Qualitätsindex Mathematik nach der Intervention nützlicher, Teilnehmende mit besonders niedrigen Indexwerten beurteilten die Mathematik jedoch nach der Intervention als noch nutzloser als zuvor.

Abschließend werden die Ergebnisse der drei empirischen Studien zusammengefasst und vor dem Hintergrund des aktuellen Forschungsstands zum Thema Schülermotivation im Klassenzimmer diskutiert. Schlussfolgerungen für zukünftige Forschung sowie für Bildungspolitik und -praxis werden gezogen.

# CONTENT

# INTRODUCTION AND THEORETICAL FRAMEWORK

## General introduction

> **Alle Pädagogen sind sich darin einig:**
> **Man muss vor allem tüchtig Mathematik treiben,**
> **weil ihre Kenntnis fürs praktische Leben größten**
> **direkten Nutzen gewährt.**
>
> **(Klein & Schimmack, 1907, p. 75)**

All teachers agree on this: rigorously studying mathematics is essential because mathematical skills are highly useful in everyday life. Indeed, in education systems around the world students learn mathematics as a core subject at school from early on: It is the second most important subject, after language arts (i.e., reading and writing), in terms of compulsory instruction time from preschool to graduation (OECD, 2016a). On average, primary school students spend 15 % of their compulsory instruction time on mathematics; in lower secondary education, mathematics makes up 12 % of the compulsory instruction time, added by an additional 11 % for natural sciences. This amount of time spent on mathematics and science instruction allows students to develop systematic and critical thinking skills and problem-solving skills—essential elements of general education (Heymann, 2013). Mathematical skills are needed to be able to evaluate scientific phenomena and innovations in 21st century life in relation to society and possible societal changes (OECD, 2016b). In everyday life, mathematical skills are used, for example, to plan one's personal budget, to tip appropriately, and to understand and interpret statistics. Furthermore, mathematical skills are required in many higher level studies and professions in a wide variety of fields ranging from engineering and computer sciences to social work and education (e.g., Joint Economic Committee, 2014). In short, mathematical skills are needed to empower individuals to function successfully in our world (Wentzel & Brophy, 2014).

Despite the obvious importance of mathematical skills in everyday life, empirical studies have revealed a huge variation across countries in 15-year-olds' beliefs about the usefulness of mathematical skills (e.g., Reiss, Sälzer, Schiepe-Tiska, Klieme, & Köller, 2016). Moreover, researchers have found a downward trend throughout elementary school and secondary school in the usefulness students attribute to mathematical skills in Western countries including Canada (e.g., Chouinard & Roy, 2008), the United States (e.g., Wigfield et al., 1997), Australia (e.g., Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002), and Germany (e.g., Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2017). At the age of 15, students in Germany consider mathematical skills to be less useful for future education and occupations than students in other OECD countries (i.e., below OECD average, Reiss et al., 2016). In addition, students in Germany do not see the applicability of secondary school mathematics in daily life (Heymann, 2013). This negative attitude toward mathematics also can be seen in their academic choices: Enrolment rates in courses and study programs related to science, technology, engineering, and mathematics (STEM) remain low (e.g., Acatech & Körber, 2015; Reiss et al., 2016).

Students' beliefs about the usefulness of mathematical skills influence not only their decision to enroll in courses, but also their academic behavior and achievement. In empirical studies framed within expectancy-value theory (EVT) of achievement motivation investigation has been made into the importance of students' beliefs about the usefulness of learning a topic, which is referred to as utility value, for their current and future personal goals (Eccles et al., 1983). In these studies, utility value as well as competence beliefs and other value beliefs correlated with learning-related behavior such as effort, perseverance, and cognitive engagement, and scholastic achievement in mathematics (for summaries of research findings, see e.g., Roeser, Eccles, & Sameroff, 2000; Vansteenkiste et al., 2004; Wigfield, Tonks, & Klauda, 2009). In short, if students see the relevance of the mathematical topics being learned in class, they might have more success learning mathematics at school.

The aim of this dissertation is to determine how the usefulness of mathematical skills be can conveyed effectively in the classroom and whether doing so enhances students' motivation to learn mathematics. In line with theoretical assumptions made in EVT (Eccles et al., 1983), two approaches are taken to convey the relevance of learning about mathematics topics in the classroom: through everyday instructional practices and through targeted scientific interventions.

In EVT students' socializers in the classroom, namely teachers and peers, play a key role in the development of students' values (Eccles et al., 1983). In numerous studies investigation has been made into the effectiveness of various instructional approaches for fostering students' motivation and achievement (for a review, see e.g., Fredricks, Blumenfeld, & Paris, 2004). However, the influence of relevance-oriented instructional methods and the role of classmates' attitudes toward learning mathematics on the development of students' values ostensibly have not been investigated. Results of studies conducted in the United States of the potential of relevance interventions in the classroom indicate that the value students attribute to scholastic learning can be raised through rather short and simple targeted writing activities (e.g., Hulleman & Harackiewicz, 2009). However, little is known about the effectiveness of various intervention approaches to conveying the relevance of mathematics, the breadth and sustainability of the intervention effects, and the processes through which scientific relevance interventions work.

For this dissertation, three empirical studies are conducted to investigate the educational potential of relevance orientation in genuine classroom settings and relevance interventions in mathematics class. In Study 1 investigation is made into whether relevance-oriented teaching practices and perceived classmates' mathematics-related value beliefs are associated with students' mathematics-related value beliefs. In Study 2 the effects of two relevance interventions conducted by researchers in the classroom on students' competence beliefs, effort, and achievement in mathematics are assessed. In Study 3, the processes underlying the effects of the two relevance interventions are explored by investigating the predictors of students' responsiveness to the relevance interventions and the effect it has on their utility value beliefs.

In **Chapters 1 to 3**, the broad research context and theoretical framework of the three empirical studies conducted are described and the research questions explored in this dissertation are presented. In **Chapter 1**, the expectancy-value model of achievement motivation and its central constructs are presented and compared with different theories of academic motivation. Research findings concerning the development and educational relevance of students' utility value beliefs, competence beliefs, and other value beliefs are summarized. In **Chapter 2**, various ways teachers can foster students' motivation in the classroom through everyday educational practices and peer influences on the development of students' motivation are described. Furthermore, motivational interventions that enhance students' motivational beliefs in classroom settings are outlined and current challenges regarding the design, implementation, and evaluation of relevance interventions in classroom contexts are discussed. In **Chapter 3**, the research questions explored in the dissertation are presented. In **Chapters 4 to 6**, the three empirical studies are presented. In the final part of this dissertation (**Chapter 7**), the findings of the three empirical studies are summarized and then discussed with regard to the theoretical framework and the implications for future research and educational practice.

# 1

# The expectancy-value model of achievement motivation

## 1.1 Background and definition

For decades, researchers and educational practitioners have tried to understand how students' achievement motivation develops, why individuals differ in motivation, and which contextual factors affect motivation. In education, achievement motivation can be defined as the energy students bring to academic tasks or subjects, the beliefs, values, and goals that determine which tasks or courses they pursue, and their persistence in achieving those goals (Wentzel & Wigfield, 2009). Expectancy-value frameworks have been particularly influential in achievement motivation research, with early versions dating back to the 1950s. In line with findings from a series of laboratory studies, Atkinson (1957) postulated that expectancies for success and incentive values determine achievement-related behaviors such as task choice and task persistence. He defined expectancy of success as a person's belief about the probability of succeeding on a given task and defined incentive value as a person's relative desirability to succeed on a given task. These two constructs of motivation were assumed to be inversely related: Atkinson (1957) argued that individuals value most the kinds of task they believe are difficult to do.

In the 1980s, Eccles et al. (1983) transferred Atkinson's theory to achievement-related behaviors in educational contexts. Inspired by Atkinson's work (Atkinson, 1957; Atkinson & Feather, 1966), research on achievement values (Battle, 1965, 1966) and intrinsic and extrinsic motivation (Deci, 1975; Deci & Ryan, 1985), as well as their own studies conducted in real-world achievement situations, the authors presented their own version of EVT. They suggested expectancies and subjective task values were the most central elements determining students' motivation-related behaviors (e.g., choice of and engagement and persistence in an activity) as well as performance. Eccles et al. (1983) refined and broadened the early expectancy-value approaches by differentiating two kinds of expectancy beliefs, namely those about success and ability, as well as four components of task value beliefs, namely intrinsic value, attainment value, utility value, and relative cost. Furthermore, they defined a range of social-psychological and cultural factors assumed to predict students' expectancy beliefs and value beliefs.

In the Eccles et al.'s (1983) definition, expectancies for success refers to how successful a person thinks he or she will be in the future (in a domain or on a given task), whereas ability beliefs refers to an individual's perceived current competence in a domain. Together with perceived task difficulty, students' ability beliefs are thought to influence task-specific expectancies for success. Intrinsic value (also called interest value or interest-enjoyment value) is defined as the enjoyment an individual obtains from doing a task or activity. Attainment value refers to the importance of doing well on a task or activity for an individual, as well as the individual's perception of a task to be central to his or her personal identity. Utility value reflects how useful a task or activity is perceived to be for an individual's current or future plans. Relative cost is defined as an individual's perception of the effort needed to accomplish a task and as the negative impact of task engagement on other valued activities (Eccles & Wigfield, 2002).

Since its development, Eccles et al.'s (1983) renowned EVT has provided a framework for investigation into the development, antecedents, effects, and promotion of student motivation in genuine educational settings. Incorporating various personal, affective, and environmental influences with students' cognitions (i.e., their expectancy beliefs and value beliefs), Eccles et al.'s theory can be placed in a social cognitive tradition of research on human motivation and achievement (e.g., Bandura, 1986). Social cognitive approaches assume that, embedded in a social environment and influenced by relationships with the activity and significant others, personal cognitions (e.g., beliefs about the self and self-regulatory processes) determine how a person acts in his or her social context (Pintrich & Schrauben, 1992).

Eccles et al.'s (1983) expectancy-value model is depicted in Figure 1. In this model, expectancy beliefs and value beliefs are the most direct predictors of students' achievement-related choices, behaviors, and performance. In contrast to Atkinson's (1957) assumption, expectancy beliefs and value beliefs are postulated to correlate positively: Individuals are assumed to value most the tasks they believe they are good at. In their model, Eccles et al. (1983) furthermore expect that individual factors such as personal goals, dispositions, and achievement experiences are related to expectancy beliefs and value beliefs. In particular, the beliefs and behaviors of students' socializers such as parents, teachers, and peers, as well as students' perceptions thereof are assumed to play a role in the development of expectancy beliefs and value beliefs. Lastly, the cultural milieu in which a student grows up is assumed to contribute to his or her expectancy beliefs and value beliefs.

### 1.2 EVT in the context of other theories of motivation

The central constructs defined in EVT can be linked to basic questions about student motivation. Students' expectancy beliefs concern students' answers to the questions: Can I do this task? If I succeed or fail, why did it happen? Students' value beliefs concern their answers to the questions: Do I want to do this task? Is it important and enjoyable to do? (e.g., Wentzel & Brophy, 2014). Students' answers to these questions are central to their subsequent behavior when approaching a task. In addition to EVT (Eccles et al., 1983), other theories of motivation such as attribution theory, implicit theories of intelligence, self-efficacy theory, self-determination theory, and interest theory address these critical questions concerning motivation—mainly with a focus on either expectancies or values. Accordingly, a range of conceptually similar but theoretically different constructs of motivation coexist.

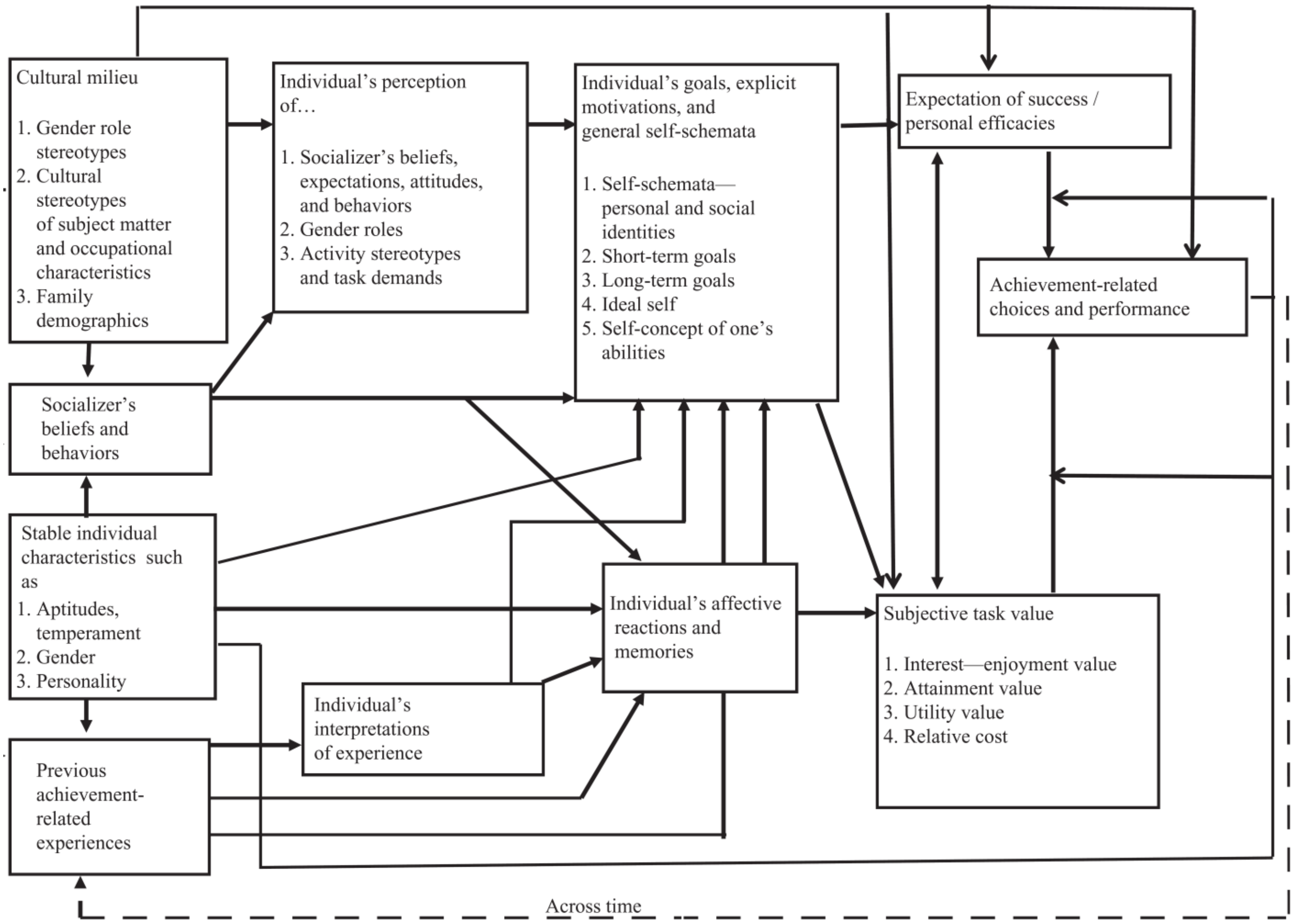



**Figure 1:** The modern expectancy-value model of achievement-related choices, behaviors, and performance (Eccles, 2011).

Unlike other theories of motivation, EVT (Eccles et al., 1983) integrates students' expectancies and values as central motivation-related beliefs in one model, links them to outcomes, and defines their antecedents. Through its comprehensiveness, EVT provides a sound frame-work for investigation into students' development of achievement motivation in genuine learning settings. Because findings related to expectancy- and value-related constructs stemming from other theories of motivation influenced the conceptualization of EVT and may complement research on EVT, the following section provides a brief overview of the most important theories of motivation related to EVT. Expectancy-related constructs such as self-efficacy beliefs and beliefs about intelligence, which have some overlap with expectancies for success and ability beliefs, are presented. Due to their conceptual similarities and empirical overlaps, expectancy-related constructs are subsumed under the notion of *competence beliefs* (cf., Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006; see also **1.2.1**). In addition, constructs of motivation similar to value beliefs such as intrinsic motivation, extrinsic motivation, and interest are briefly explained. Each section closes with a summary of the theoretical and empirical differences and overlaps of competence- and value-related motivational constructs and implications for this dissertation.

### 1.2.1 Competence-related theories and constructs of motivation

**Attribution theory and theories of intelligence**

Rooted in an expectancy-value tradition, attribution theory (1986, 2006) assumes that students' achievement motivation is determined by their interpretation of achievement outcomes. *Achievement attributions*, that is, the perceived causes of success or failure on a task (e.g., ability, effort, task difficulty, etc.), constitute the core constructs of attribution theory. These achievement attributions can be classified according to the three dimensions stability, locus, and controllability, which are believed to influence achievement motivation and expectancy beliefs. In his work, Weiner (1979, 1985) found that if a student attributed an outcome to stable causes such as low aptitude, this had a greater influence on students' subsequent expectancies for success on a task than attributing the success or failure to unstable causes such as effort. The assumption that students' perception of their abilities impacts on their expectancies for success influenced Eccles' EVT.

Building on Weiner's (1986) attribution theory, implicit theories of intelligence (Dweck, 1999; Dweck & Leggett, 1988; Dweck & Molden, 2005) suggest that individuals have different theories about the nature of intelligence in terms of stability and control. More precisely, students who believe that intelligence is basically fixed and unchangeable are believed to have an *entity theory* of intelligence; students who think that intelligence is malleable and can be developed consciously are believed to hold an *incremental theory* of intelligence. These implicit theories of intelligence are believed to influence students' achievement goals as well as their behavior when faced with academic challenges. Entity theorists are believed to set performance goals (i.e., they seek favorable judgment and avoid negative judgment concerning their competence) and to reduce or withdraw effort when they think their aptitude is too low to complete an academic task successfully. In contrast, incremental theorists are believed to set mastery goals (i.e., they seek to increase their competence and to master the learning material) and to see usefulness in making effort to overcome academic challenges (Dweck, 1986; see also learning goal theory, e.g., Ames, 1992; Dweck & Leggett, 1988).

**Self-efficacy theory**

Self-efficacy theory (Bandura, 1977a, 1997) is rooted in a social cognitive approach to understanding motivation (Bandura, 1986) and is somewhat similar to the concepts of expectancies for success and ability beliefs in EVT. More precisely, self-efficacy theory focuses on students' beliefs about the behaviors needed to complete a given task successfully (i.e., *outcome expectations*) and students' perceived capability to perform these behaviors (i.e., *efficacy expectations* or *self-efficacy*). In genuine educational contexts, these two constructs correlate positively: Students who are confident about their skills often expect good grades on their examinations (Bandura, 1997). According to social cognitive theory, there is a reciprocal relationship between students' social-cognitive environment and their self-efficacy. Students' self-efficacy is assumed to influence their learning environments (e.g., finding study partners) and learning behaviors (e.g., using effective learning strategies). Vice versa, students' learning environments (e.g., their teachers' or peers' behaviors) and the outcomes of their learning behaviors (e.g., achievement) are assumed to have an impact on students' subsequent self-efficacy.

Compared to the expectancy-related constructs defined in EVT, self-efficacy differs from expectancies for success theoretically in that it is not about a student's beliefs about his or her probability of completing a task successfully, but rather the probability of being able to perform the behaviors needed to accomplish the task. It also differs from ability beliefs, also referred to as self-concept of abilities: Self-efficacy is relatively context-specific judgments about one's competence, which are malleable because of their task dependence and are generally future-oriented. In contrast, ability beliefs/self-concept of abilities are hierarchically structured and consist of several domain-specific self-perceptions of competence (e.g., in the academic, physical, and social domains), which are past-oriented and more stable because of their sense of generality (Schunk & Pajares, 2009). However, despite these theoretical differences, the empirical distinction between expectancies for success, self-concept of abilities, and self-efficacy is not straightforward. In fact, when measured at the task level, expectancies for success and self-efficacy often are operationalized in a similar manner (e.g., Wigfield & Eccles, 2002). At the domain level, factor analyses have shown that students' expectancies for success load on the same factor as academic self-concept (e.g., Eccles & Wigfield, 1995), and self-efficacy—which also has been investigated at the domain level—also shows considerable empirical overlaps with academic self-concept (see Bong & Skaalvik, 2003, for a detailed discussion of the distinctiveness of self-efficacy and self-concept; Hughes, Galbraith, & White, 2011; Pietsch, Walker, & Chapman, 2003). Research on the empirical overlaps between domain-specific self-efficacy and expectancies for success is scarce, but findings seem to indicate empirical distinctiveness at least when self-efficacy is measured according to present competence and expectancies for success are operationalized according to future competence (e.g., Yong, 2010).

**Summary and implications for the present research**

Similar to EVT, both attribution theory and theories of intelligence acknowledge the importance of competence-related beliefs for students' academic behavior and achievement. However, they differ from Eccles et al.'s (1983) EVT in that they postulate that achievement strivings are determined by students' beliefs about the stability and controllability of their abilities—rather than by ability beliefs as motivational dispositions. Findings concerning attribution theory have influenced EVT and the assumption that students' domain-specific ability beliefs influence students' expectancies for success.

The concept of self-efficacy as framed within self-efficacy theory is similar to students' expectancies for success in that they both focus on future-oriented, task-specific competence beliefs. Their theoretical differences rely on nuances: Expectancies for success are related to the probability of success on a task, whereas self-efficacy is related to the probability of performing certain behaviors when working on a task. On an empirical level, however, students' self-efficacy, expectancies for success, and self-concept of abilities are undistinguishable, unless operationalized at different levels (domain vs. task). Thus, the current dissertation will distinguish between domain- and task-specific competence beliefs while equally including research on self-efficacy, self-concept of abilities, and expectancy beliefs.

### 1.2.2 Value-related theories and constructs of motivation

#### Intrinsic and extrinsic motivation and self-determination theory

Eccles et al.'s (1983) development of the four value beliefs in EVT was influenced by previous work on intrinsic and extrinsic motivation (Deci, 1975; Deci & Ryan, 1985; Ryan & Deci, 2000), which also constitutes the basis of self-determination theory (SDT). *Intrinsic motivation* is considered a human disposition and is defined as "the doing of an activity for its inherent satisfaction rather than for some separable reasons" (Ryan & Deci, 2000, p. 56). Intrinsically motivated activities such as play and active learning take place independently from external incentives, which is why they are considered to be autonomous. In contrast, *extrinsic motivation* refers to activities that are "done in order to attain some separable outcome" (Ryan & Deci, 2000, p. 60). Although intrinsic motivation and extrinsic motivation are antagonistic constructs, humans also have a disposition to assimilate and internalize nonintrinsically motivated practices and values from their social and cultural environment. Through processes of internalization and integration, originally extrinsic activities can become increasingly autonomous and, consequently, intrinsically motivated (Deci & Ryan, 1985; Ryan & Deci, 2000). Accordingly, SDT postulates that there is a continuum from extrinsic motivation to intrinsic motivation, comprising various types of extrinsic motivation in which the regulation of academic behavior becomes increasingly internalized, integrated, and self-determined.

Inspired by theories of intrinsic, extrinsic, and self-determined motivation, reasons to engage in an activity which pertain to an individual's self, personal interests, identities, and goals were introduced to EVT in the form of value beliefs. Yet in contrast to SDT, which focuses on antagonistic *types* of motivation (intrinsic vs. extrinsic), EVT considers different value beliefs and thus different *degrees* of intrinsic or extrinsic motivation to contribute to a comprehensive perception of value (Wigfield et al., 2009). External motives to engage in a task are not per se considered undesirable as long as they are personally important and thus contribute to overall high degrees of subjective task value. Accordingly, both intrinsic and extrinsic aspects of motivation can be found to varying degrees in the concepts of intrinsic value, attainment value, and utility value (Eccles, 2005). A person who intrinsically values a task does not do it as a means to another end but rather because he or she enjoys it, which comes very close to Ryan and Deci's (2000) definition of intrinsic motivation. If an individual perceives high levels of attainment value, both the activity and its outcomes are important to the person's self and identity, which comprises intrinsic but also extrinsic motives. Similarly, a person who thinks that a task is useful for his or her future plans acts to achieve a personally valued outcome. In both cases, the behavior is highly internally regulated, personally important, and more (attainment value) or less (utility value) self-determined. This implies that attainment and utility values have ties to both intrinsic and extrinsic motivation. There are no explicit overlaps between EVT and externally regulated types of extrinsic motivation in which engaging in a task relies on an individual's desire to receive a reward or avoid punishment.

### Interest theory

Interest theory (e.g., Schiefele, 1991; 2009) is concerned with answers to the question "Why do I want to do this task?". The terms interest and motivation often are used synonymously in everyday conversation. However, in research in education, motivation refers to an individual's drive to do a certain activity in a specific situation and interest actually represents a possible determinant of motivation. Interest is related to either a specific task or domain and, accordingly, comprises two conceptions: *situational interest*, a temporary, task- and situation-specific psychological state characterized by high attention and positive emotions; and *individual interest*, an enduring personal orientation toward a domain—a stable character trait. Situational interest is assumed to influence students' intrinsic motivation to learn directly; individual interest is assumed to impact intrinsically motivated learning either directly or indirectly via the experience of situational interest (Schiefele, 2009). Under certain conditions, situational interest can become individual interest (Hidi & Harackiewicz, 2000). According to Schiefele (1991; 2009), individual interest also can be defined as a relatively stable set of valence beliefs. These valences refer to a) feelings associated by an individual with a domain of interest and to b) the personal value and significance attributed by an individual to a domain of interest. In this definition, the feeling-related aspect of individual interest has a considerable conceptual overlap with intrinsic task value found in EVT. In addition, the value-related aspect of individual interest shares certain aspects with attainment and utility value, namely that a domain of interest has personal importance and relates to personal goals (Schiefele, 2009; Trautwein et al., 2013). Furthermore, interest and value beliefs are both constructs which vary in degree.

Despite some similarities, the theoretical complexity of individual and situational interest is not fully captured in the definition of intrinsic value (Wigfield et al., 2009). Nevertheless, similar to competence-related constructs, the empirical distinctiveness of the constructs of interest, intrinsic value, and intrinsic motivation depends on their distinct operationalization, in particular, their measurement level. At the domain level, individual interest, intrinsic value, and intrinsic motivation often have been measured using similar items, and overlaps in wording also may occur with items used to measure attainment value (Wigfield & Cambria, 2010). In contrast, when measured at different levels, individual (i.e., domain-specific) interest and task-related intrinsic value beliefs have been shown to be empirically distinct constructs (Hulleman, Durik, Schweigert, & Harackiewicz, 2008).

### Future time perspective

Future time perspective (FTP) (Nuttin, 2014; Nuttin & Lens, 1985; Volder & Lens, 1982) is a cognitive theory of motivation concerning present anticipation of future goals and beliefs about the instrumentality of engaging in tasks to attain these personal goals (i.e., *instrumentality beliefs*). In contrast to learning goal theory, which is concerned with approaches towards learning (performance vs. mastery as learning goals, e.g., Ames, 1992; Dweck & Leggett, 1988), FTP focuses on personal life goals and the time perspective taken on these goals. Accordingly, goals are differentiated into immediate, present goals and more long-term, future goals. In addition, personal goals may refer to different areas of life such as school, career, social relationships, personal development, and leisure (Peetsma, 2000). A FTP is created by setting highly valued goals in the rather distant future and by engaging in long-range projects to achieve these goals (Volder & Lens, 1982). Accordingly, FTP combines cognitive aspects of motivation (ability to relate present behavior to distant goals) with dynamic aspects of motivation (ascribing relevance to personal goals). Individuals who take on a FTP are assumed to perceive their current behavior as more instrumental to accomplish both short- and long-term goals and, thus, to value task engagement more than people who are focused on immediate goals (Simons, Vansteenkiste, Lens, & Lacante, 2004). Recent work on the instrumentality perspective has incorporated aspects of EVT and SDT into the definition of four types of instrumentality beliefs based on the degree of utility value (low vs. high) and on the type of regulation (internal vs. external) (Simons et al., 2004).

Whereas utility value beliefs refers to the perceived utility of engaging in a task for attaining personal goals in general, instrumentality beliefs are more specific regarding their time perspective and the nature of the goals. In empirical studies, instrumentality beliefs have been measured with items that contain a time marker (present vs. future) and refer to a broad range of personal goals in various areas of life (Husman, Derryberry, Crowson, & Lomax, 2004). In contrast, the operationalization of utility value has been diverse, either without any time and goal specification (e.g., Eccles, Wigfield, Harold, & Blumenfeld, 1993) or with mixed time orientations referring to different goals (e.g., concerning present leisure time and future employment, Eccles & Wigfield, 1995), resulting in overlaps with items measuring instrumentality. However, value scales more recently developed in the context of EVT provide a more comprehensive and refined measurement of utility value beliefs, including time markers and goals in various areas of life (Gaspard, Dicke, Flunger, Schreier et al., 2015; see **1.3.1**).

**Summary and implications for the present research**

Theories of intrinsic and extrinsic motivation—and SDT—influenced Eccles et al.'s (1983) development of value beliefs in EVT, some of which are more (intrinsic value) or less (utility value) intrinsically motivated. The major difference between SDT and EVT is that intrinsic motivation and extrinsic motivation constitute oppositional types or qualities of motivation. In contrast, EVT defines value beliefs (intrinsic, attainment, utility) that vary in quantity and, when in greater amounts, contribute to an overall higher level of motivation.

Interest theory distinguishes between situational interest as a temporary psychological state of enjoying a task and individual interest as an enduring character trait shaped by personal value attributed to a domain of interest. Although the conceptual complexity of the constructs of interest is not represented in the definition of intrinsic value, interest and intrinsic value refer to similar basic concepts (Wigfield et al., 2009). Empirically, interest, intrinsic value, and intrinsic motivation often have been operationalized with similar items, which is why research on students' interest and intrinsic motivation also are consulted when referring to empirical findings concerning students' intrinsic value beliefs in this dissertation.

Similar to utility value beliefs as defined in EVT, the core construct of FTP—instrumentality beliefs—also deals with how important task engagement is perceived to be in order to attain personal goals. Theoretically, instrumentality beliefs rely on the time distinction between broad ranges of present or future goals, whereas utility value beliefs relate to general future goals (Eccles et al., 1983). However, the empirical operationalization of utility value beliefs has been mixed (i.e., with or without time markers, referring to general utility or specifying a goal), resulting in notable empirical overlaps with items measuring instrumentality. Taking a broad view on the utility value construct of motivation, including utility for present and future personal goals, the notion of students' *relevance perceptions* will subsume several utility-related constructs and respective research in this dissertation (cf., Priniski, Hecht, & Harackiewicz, 2017). In addition to studies of instrumentality beliefs (e.g., Husman et al., 2004), studies in which terms such as "curricular meaningfulness" (e.g., Roeser et al., 2000), "instrumental motivation" (e.g., OECD, 2016b), and "functional relevance (of learning)" (e.g., Woolley, Rose, Orthner, Akos, & Jones-Sanpei, 2013) were used also are taken into account when reviewing research on students' relevance perceptions.

## 1.3 Educational relevance of competence beliefs and value beliefs

### 1.3.1 Measurement of competence beliefs and value beliefs

The previous section on the theoretical and empirical distinction of competence- and value-related constructs in theories of motivation highlighted an important issue in empirical research in education: the correspondence between the notional definition of a construct and its empirical measurement. If the empirical operationalization of a construct (e.g., the wording of items on a questionnaire and the combination of various items into a scale) is inconsistent with the respective theory and does not capture all facets of its notional definition, conceptual confusion and empirical overlap may ensue (e.g., Usher, 2016; Wigfield et al., 2009). Before presenting empirical findings concerning the educational relevance of competence beliefs and value beliefs, a brief outline is given of important aspects and current research strands concerning the measurement of competence beliefs and value beliefs as defined in EVT. Conclusions for the operationalization of the constructs of motivation investigated in this dissertation are drawn.

As mentioned before, expectancies for success can be measured at the task level or domain level. At the task level, expectancies for success often have been operationalized in a similar way with self-efficacy referring to students' beliefs about whether they can complete a specific task such as a homework assignment correctly (Wigfield & Eccles, 2002). As expectancies for success have been found to be empirically indistinguishable from self-concept of abilities at the domain level (Eccles, Wigfield et al., 1993; Eccles & Wigfield, 1995), students' self-concept within a certain school subject frequently has been used as the domain-specific expectancy belief in studies framed in EVT. Results of recent research indicate that domain-specific self-concept and task-specific self-efficacy constitute empirically distinct factors in science (Jansen, Scherer, & Schroeders, 2015). As a result, in the empirical studies presented in this dissertation, students' self-concept in mathematics will be used to represent the domain-specific aspect of students' competence beliefs, while students' self-efficacy concerning homework in mathematics will represent the task-specific aspect of students' competence beliefs.

Many instruments have been developed to assess students' value beliefs as defined in EVT (e.g., Eccles et al., 1983; Eccles, Wigfield, et al., 1993; Wigfield & Eccles, 2000). However, in recent research notable inconsistency in the use of those measures has been criticized (Gaspard, Dicke, Flunger, Schreier, et al., 2015; Trautwein et al., 2013). In particular, value beliefs have not always been captured with all of their components. First, the cost component frequently has been neglected (e.g., Eccles, Wigfield et al., 1993; Jacobs et al., 2002; Wigfield & Eccles, 2000). Second, positive value beliefs (intrinsic, attainment, and utility values) often have been overlooked or combined into simpler scales: In many studies a small number of value items have been used to create a single and rather general value scale (e.g., Eccles, Wigfield et al., 1993; Wang, 2012). In other studies one positive value belief has been assessed separately in addition to a combination of the two other positive value beliefs, for instance, intrinsic value and a combination of attainment/utility value belief ("importance value", e.g., Durik, Vida, & Eccles, 2006; Watt et al., 2012) or interest/attainment value belief and utility value (e.g., Battle & Wigfield, 2003).

In recent empirical research separate measures have been used to assess the four distinct components of value (intrinsic, attainment, utility, and cost) and to increase the explanatory power of value beliefs (Conley, 2012; Trautwein et al., 2012). Further, cost has been treated separately from positive value beliefs in an expectancy-value-cost model of achievement motivation (Barron & Hulleman, 2015). Beyond that, Trautwein et al. (2013) and Gaspard, Dicke, Flunger, Schreier, et al. (2015) argued that the notional richness of value beliefs was not fully covered in the measures used—even when the four value components were measured separately. Consequently, Gaspard, Dicke, Flunger, Schreier, et al. (2015) suggested a refined theoretical differentiation of three of the four value components into several subfacets, which may allow conceptual confusion and nominal confusion extant in prior research to be cleared up. Multiple subfacets of utility value have been assessed while taking into account research on life goals as reported in studies framed in a FTP (e.g., Simons et al., 2004). Analyzing the data under investigation in this dissertation, Gaspard, Dicke, Flunger, Schreier, et al. (2015) showed the empirical distinctiveness of ten subcomponents of value (see Table 1). This refined instrument by Gaspard, Dicke, Flunger, Schreier, et al. (2015) is used in this dissertation to investigate students' value beliefs.

Table 1
*Facets of value beliefs according to Gaspard, Dicke, Flunger, Schreier, et al. (2015)*

| 1st order factor: Value component | 2nd order factors: Value subfacets | Sample item |
|---|---|---|
| Intrinsic value | -- | Math is fun to me. |
| Attainment value | Importance of achievement | It is important to me to be good at math. |
| | Personal importance | Math is very important to me personally. |
| Utility value | Utility for school | Being good at math pays off, because it is simply needed at school. |
| | Utility for daily life | Math is directly applicable in everyday life. |
| | Social utility | I can impress others with intimate knowledge in math. |
| | Utility for job | Good grades in math can be of great value to me later on. |
| | General utility for future life | I will often need math in my life. |
| Cost | Effort | Doing math is exhausting to me. |
| | Emotional cost | When I deal with math, I get annoyed. |
| | Opportunity cost | I have to give up a lot to do well in math. |

### 1.3.2 Relationship between students' competence beliefs and value beliefs

In addition to investigating effective ways to measure students' competence beliefs and value beliefs, researchers adhering to EVT have long explored the relationship between students' competence beliefs and value beliefs: Do students value tasks they find challenging (i.e., negative relationship)? Or do they value tasks they find easy (i.e., positive relationship)? Students' competence beliefs and value beliefs both have been found to be highly domain-specific (Marsh, 1990; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006; Trautwein, Lüdtke, Marsh, & Nagy, 2009), which means that they may differ intraindividually *across* achievement domains: A student may think he or she is competent in a subject such as English but not in another domain such as mathematics, and this is the same with students' value beliefs. Accordingly, students' competence beliefs and value beliefs can be entirely unrelated across domains. *Within* a domain, competence beliefs and value beliefs have been found to form distinctively as early as Grade 1 (Eccles, Wigfield et al., 1993), and to correlate positively (with particularly high intercorrelations between intrinsic value and competence beliefs; e.g., Bong, 2001b; Eccles, Vida, & Barber, 2004; Nagengast et al., 2011). This positive relationship between students' competence beliefs and value beliefs also has been found to persist and even increase as children grow older (Jacobs et al., 2002; Trautwein et al., 2012; Wigfield et al., 1997).

These results indicate that students value the tasks they believe they can complete successfully and the subjects they believe they can succeed in. However, the direction of influence is not totally clear: Students might either learn to develop their ability to perform highly valued activities or start valuing the tasks they are already good at doing (cf., Wigfield et al., 2009). Empirical evidence concerning the direction of influence is scarce. In a longitudinal study in which students were surveyed from Grade 1 through Grade 12, a strong association was found between changes in students' competence beliefs and the development of their task values (Jacobs et al., 2002). In addition, changes in competence beliefs were found to reduce the intraindividual variance in the development of task values by at least 40 % in mathematics, language arts, and sports. The authors concluded that students value subjects *because* they are good at them but conceded that their correlational study design did not exclude influences in the opposite direction, too. Furthermore, in a study of the reciprocal effects of interest and self-concept in secondary school mathematics, Marsh, Trautwein, Lüdtke, Köller, and Baumert (2005) found a significant effect of self-concept in mathematics on subsequent interest in mathematics. In contrast, prior interest in mathematics had only a small, marginally significant effect on subsequent self-concept in mathematics, which again supports the assumption that students value subjects they are good at rather than the other way around. However, as Jacobs et al. (2002) acknowledged, correlational studies are limited in shedding light on cause and effect (cf., Shadish, Cook, & Campbell, 2002). By manipulating students' value beliefs and investigating their effects on students' competence beliefs, experimental studies could help to clarify if there is an influence in the opposite direction—hitherto a research gap.

### 1.3.3 Relationship between competence beliefs and value beliefs with academic outcomes

The empirical validity of the Eccles et al. (1983) expectancy-value model has been tested in numerous real-life achievement contexts in education. The original goal of the model was to understand gender differences in education- and occupation-related choices in mathematics (Eccles et al., 1983; Eccles, 1994). Over the past three decades, the model has been applied to numerous other achievement-related settings in education. Despite its comprehensiveness, the focus of most studies grounded in Eccles et al.'s (1983) EVT has been the right part of the model depicted in Figure 1, that is, the extent and development of students' competence beliefs and value beliefs and the influence of those beliefs on education-related choices, behavior, and achievement. Overall, students' competence beliefs and value beliefs have been found to predict several positive learning outcomes such as effort, persistence, task engagement, course enrollment, and achievement in various subjects (for recent summaries of research findings, see e.g., Barron & Hulleman, 2015; Hulleman, Barron, Kosovich, & Lazowski, 2016; Usher, 2016; Wigfield et al., 2009).

More precisely, competence beliefs, including expectancy for success, academic self-concept, and academic self-efficacy, have been found to predict academic achievement in various subjects (e.g., Bong, 2001a; Denissen, Zarrett, & Eccles, 2007; Jansen et al., 2015; Marsh & Craven, 2006; Roeser et al., 2000; Simpkins, Davis-Kean, & Eccles, 2006; meta-analysis by Valentine, DuBois, & Cooper, 2004). In many cases, competence beliefs remain strong predictors of achievement even when controlling for prior achievement, highlighting the importance of personal beliefs to influence future performance (e.g., Jacobs et al., 2002). Positive associations also have been found with choice of activity, course, or career (e.g., Durik et al., 2006; Eccles et al., 1983; Jansen et al., 2015; Nagengast et al., 2011; Simpkins et al., 2006). In further research students' competence beliefs predicted academic effort (e.g., Trautwein & Lüdtke, 2009; Trautwein, Lüdtke, Kastens, & Köller, 2006), the use of successful metacognitive learning strategies (Zimmerman & Schunk, 2011), and cognitive engagement (Walker, Greene, & Mansell, 2006). In addition, students with more positive competence beliefs have had fewer reported behavioral problems at school (Roeser et al., 2000) and lower levels of negative academic emotions such as stress and depression (Bandura, 1997) than students who believe less in their academic capabilities.

Intrinsic value, attainment value, and utility value or composite measures of these positive value beliefs have been shown to predict academic performance and choices (e.g., Battle & Wigfield, 2003; Chow, Eccles, & Salmela-Aro, 2012; Durik et al., 2006; Hulleman et al., 2008; Nagengast et al., 2011; Roeser et al., 2000; Simpkins et al., 2006), effort (e.g., Cole, Bergin, & Whittaker, 2008; Trautwein et al., 2006; Trautwein & Lüdtke, 2009), and cognitive engagement (e.g., Walker et al., 2006; Walker & Greene, 2009). Furthermore, students with positive intrinsic, attainment, and utility value beliefs have had fewer reported behavioral problems at school, better peer relationships, and less emotional distress at school than their peers with more negative value beliefs (Roeser et al., 2000). While controlling for other value beliefs, Battle and Wigfield (2003) found perceived psychological cost of attending graduate school negatively predicted college students' intentions to enter graduate school. Conducting latent profile analyses, Conley (2012) found that students with high-cost profiles reported higher levels of negative affect and achievement than students in low-cost profiles. While effort and opportunity cost were found to predict students' intentions to leave science-related courses, psychological cost was unrelated (Perez, Cromley, & Kaplan, 2014).

Generally, students' competence beliefs most strongly predict achievement, whereas students' value beliefs most strongly predict education-related choices and decisions (e.g., Wigfield et al., 2009)—even in the long term. For example, Durik et al. (2006) found students' value beliefs in elementary school predicted educational choices at secondary school. Further research findings indicate that the relationship between students' motivation-related beliefs and academic outcomes gain strength in age (Denissen et al., 2007; Eccles et al., 1983). In addition, there seems to be a synergistic relationship between students' motivation-related beliefs: The effect of competence beliefs on education-related choices and achievement is stronger when students also reported more positive value beliefs, and vice versa (Nagengast et al., 2011; Trautwein et al., 2012).

### 1.3.4 Developmental trajectories of competence beliefs and value beliefs

The educational importance of students' competence beliefs and value beliefs notwithstanding, results of empirical studies of their developmental trajectories indicate that these motivational beliefs become increasingly negative as students progress through school—in particular, in mathematics (e.g., Frenzel, Goetz, Pekrun, & Watt, 2010; Gaspard et al., 2017; Jacobs et al., 2002; Nagy et al., 2010; Watt, 2004). The developmental trajectory of students' mathematics-related value beliefs declines steadily when using a combined score of intrinsic value, attainment value, and utility value (Jacobs et al., 2002) and when using separate measures of intrinsic value or interest, attainment value, and utility value (Frenzel et al., 2010; Gaspard et al., 2017; Watt, 2004). In some studies, the decrease in students' mathematics-related intrinsic value has been found to level off at the end of secondary school (Frenzel et al., 2010; Watt, 2004). Interestingly, it is students' utility value beliefs about mathematics, especially its usefulness for daily life and for work, which show a steep linear decrease from Grade 5 to Grade 12 (Gaspard et al., 2017). Girls report lower levels of motivational beliefs about mathematics than boys throughout all grade levels (e.g., Frenzel et al., 2010; Gaspard et al., 2017; Jacobs et al., 2002; Nagy et al., 2010; Watt, 2004).

Stage-environment fit theory (Eccles, Midgley et al., 1993) proposes a mismatch between students' needs and interests and their school environment as an explanation for students' downward trajectory in competence beliefs and value beliefs throughout school. At least two further reasons have been proposed for students' increasingly negative competence beliefs (Wigfield et al., 2009; Wigfield & Eccles, 2002): First, systematic evaluations become more and more prevalent and important as students advance in school. Children use feedback—possibly gender-stereotyped—from teachers and parents to revise their own understanding about their competences, weaknesses, and strengths in academic domains (e.g., Jacobs et al., 2002; Jacobs & Eccles, 1992). In addition, students' cognitive capacity to process and interpret such competence-related information increases with age. Consequently, students' self-assessments become more realistic and, possibly, more negative throughout schooling. Second, students start comparing themselves to their peers as soon as they begin school: They use information about other students' achievement to judge their own performance. As students grow older and enter puberty, peers become increasingly important to them and adults continually less (e.g., Kindermann, 2007). Accordingly, social comparison mechanisms become more prevalent as students go through school, causing them to calibrate their competence beliefs. For this reason, processes of social comparison have been proposed as an explanation for the increasing association between students' competence beliefs and their academic achievement (Schunk & Pajares, 2009).

Similar processes may be responsible for the downward trajectory in students' value beliefs (cf., Schiefele, 2009; Wigfield et al., 2009). From school entry onwards, students do activities in various academic domains and their experiences doing these activities can contribute to how interesting, important, or useful they find one domain in comparison to another (e.g., Krapp, 2002). They use teachers' and parents' feedback about the importance and usefulness of academic tasks to reevaluate their own value beliefs and increasingly compare their own interests and values to those of their peers (Wigfield et al., 2006). In addition, as they grow older, students gradually refine their ideas of what is important for them and accordingly adapt the value they place on a certain domain. Cultural norms and stereotypes in particular have been seen as a reason for gender differences favoring boys in competence beliefs, value beliefs, and decisions to pursue studies in science-related fields (Eccles et al., 1983; Eccles, 2005, 2011).

Over the past several decades, the downwards trajectory of students' motivation in mathematics has received increasing public attention—often for political and economic reasons rather than for educational and societal reasons (see **Chapter 2**). The following chapters explore possible mechanisms to support students' competence beliefs and value beliefs in the classroom, focusing on the particular potential of classroom contexts and scientific interventions targeting students' perceptions of the personal relevance of mathematics for daily life, education, and career pathways.

# 2

# Fostering students' competence beliefs and value beliefs in mathematics

**Die Vermittlung einer fundierten MINT-Kompetenz [ist] unverzichtbar, um das Verständnis der elementaren Vorgänge in Natur und Technik zu unterstützen und die Bewertung von wissenschaftlichen Erkenntnissen und technischen Innovationen zu ermöglichen. (…) Dieser Bildungsauftrag zielt darauf ab, Kinder und Jugendliche mit ihrer wissenschaftlich-technisch geprägten Umwelt vertraut zu machen und sie zu befähigen, gesellschaftliche Zusammenhänge und (…) Veränderungen kompetent beurteilen zu können. Dieser Bildungsauftrag gilt für alle Jugendlichen.**

**(Renn et al., 2012, as cited in Acatech & Körber, 2014, p. 6)**

In various Western countries such as the United States, Australia, and Germany, concerns about a potential lack of qualified people to work in occupations in STEM-related fields have increased public and scientific interest in students' motives for education and career choices (e.g., Acatech & Körber, 2015; Institut der deutschen Wirtschaft, 2017; Office of the Chief Scientist, 2014). Public and political discussions about how to produce qualified workers for STEM-related occupations have resulted in numerous educational initiatives aimed at attracting students—particularly females, who are underrepresented in STEM-related occupations—to enroll in STEM-related secondary school courses and university programs (for reviews, see e.g., Kärkkäinen & Vincent-Lancrin, 2013; OECD, 2014). However, motivating students to engage in learning mathematics and science should not be directed by economic motives only. According to Renn et al. (2012), promoting the development of fundamental STEM skills at school is an educational responsibility to all students for reasons pertaining to students' personal development. More precisely, schooling should familiarize all students with their environment, in which scientific and technological phenomena are ubiquitous. Students must be enabled to understand basic processes in nature and technology and to evaluate critically scientific findings and technological innovations in relation to societal developments and changes—in short, to become informed, active, and critical members of 21st-century society (see also OECD, 2016b).

Scientific and technological literacy, however, cannot be gained without fundamental knowledge of mathematics, as, for example, interpreting data and empirical evidence requires the experienced use of mathematical tools (OECD, 2016b). Numerous empirical studies cited in the previous paragraphs have underlined the importance of students' competence beliefs and value beliefs for their academic behavior, choices, and achievement in mathematics—in short, for learning mathematics. To help students develop their technological literacy and scientific literacy, practitioners and researchers in education need to find effective ways to enhance students' mathematics-related competence beliefs and value beliefs. Although high levels of motivation to learn mathematics must be maintained all the way through elementary and secondary education, this is especially important toward the end of secondary education, when students begin to choose study programs at institutions of higher education. Thus, in this chapter possible mechanisms for supporting students' motivation in secondary mathematics education are explored.

According to their EVT, Eccles et al. (1983) assume socializers' beliefs and behaviors and students' perceptions thereof are associated with students' development of values. As teachers and classmates undoubtedly are the most important socializers in the classroom context, in the following paragraphs the potential of teachers' instructional strategies and students' perception of their classmates' behavior to influence students' motivation is described. Subsequently, the potential of scientific motivational interventions implemented in classroom contexts is explored.

## 2.1 Enhancing student motivation in mathematics through common instructional practices

In numerous empirical studies ways in which teachers can help enhance and maintain students' motivation to learn mathematics and other subjects through everyday instruction have been investigated. In literature reviews a number of teaching behaviors have been identified which have been found to be crucial for students' motivation and engagement in the classroom including, for example, autonomy-supportive teaching, establishing positive relationships, providing structure and guidance, and using appropriate feedback strategies (e.g., Fredricks et al., 2004; Urdan & Schoenfelder, 2006; Wentzel & Brophy, 2014). Before discussing relevance-oriented teaching as an instructional strategy that could boost students' motivation by enhancing their subjective value beliefs (e.g., Osborne & Dillon, 2008), the most researched instructional practices and their effects on students' motivation and other important academic outcomes are presented.

### 2.1.1 Teaching to foster student motivation

The aim of **autonomy-supportive teaching** is to help students feel autonomous in their learning. It probably is one of the most researched instructional strategies in studies of motivation within classrooms (for reviews, see e.g., Assor, 2012; Patall & Hooper, 2017). Autonomy-supportive teacher behaviors such as giving students opportunities to make choices (e.g., regarding tasks to be completed during class or as homework) and minimizing control were found to predict intrinsic motivation and emotional and behavioral engagement and to avoid disaffection (e.g., Lazarides, Rohowski, Ohlemann, & Ittel, 2016; Reeve, Jang, Carrell, Jeon, & Barch, 2004; Skinner, Furrer, Marchand, & Kindermann, 2008). In particular, the strategy of offering choices was found to improve students' intrinsic motivation, competence beliefs, and positive emotions in class (e.g., Assor, Kaplan, & Roth, 2002; Patall, Cooper, & Wynn, 2010); other autonomy-supportive strategies such as taking students' perspective, using informative language, and tailoring activities to students' interests predicted students' attitude toward and perception of the usefulness of educational tasks, positive feelings, and behavioral and cognitive engagement in class (e.g., Assor et al., 2002; Jang, Reeve, & Deci, 2010; Patall et al., 2010). In contrast, autonomy-suppressing behaviors such as suppressing criticism, controlling, and intruding were found to be associated with less behavioral and cognitive engagement, poorer self-evaluations of competence, and higher emotional costs such as anxiety (e.g., Assor et al., 2002; Noels, Clement, & Pelletier, 1999).

Strategies aimed at **establishing positive relationships** within the classroom such as providing social support and opportunities for students to cooperate have not been researched only within the literature on motivation but also as one essential element of instructional quality (for reviews, see e.g., Juvonen, Espinoza, & Knifsend, 2012; Pianta, Hamre, & Allen, 2012; Wentzel, 2009). Results of numerous studies support the assumption that positive teacher-student and student-student relationships promote both motivation and learning gains, including outcomes such as valuing schoolwork, behavioral engagement, positive emotions in class, and grades (e.g., Furrer & Skinner, 2003; Goodenow & Grady, 1993; Skinner et al., 2008; Skinner & Belmont, 1993). In secondary school mathematics class, support from the teacher was found to be a major predictor of students' mathematics-related self-concept, enjoyment, and intrinsic value, and the importance they attributed to learning mathematics (e.g., Dietrich, Dicke, Kracke, & Noack, 2015; Goodenow, 1993; Kunter et al., 2013; Lazarides & Ittel, 2012). In a longitudinal study of transition to secondary school (Midgley, Feldlaufer, & Eccles, 1989), moving to mathematics classes with greater teacher support than at elementary school enhanced the intrinsic value of mathematics for students whereas moving to classes with less teacher support undermined intrinsic value and attainment and utility value beliefs related to mathematics. Furthermore, Wang (2012) found that teacher support and promotion of cooperation in mathematics in Grade 7 correlated positively with students' mathematics-related value beliefs (interest, importance, and utility) in Grade 10, mediated through students' value beliefs in Grade 7. In contrast, behaviors associated with a negative teacher-student relationship such as punishing and putting great pressure on students to perform well were found to increase emotional costs such as test anxiety in mathematics (Pekrun, 1992).

**Providing structure and guidance** can be done by giving students clear directions, stating expectations, responding consistently, and steering students' learning process. These instructional strategies have been described as competence-supportive teaching practices (e.g., Mouratidis, Vansteenkiste, Michou, & Lens, 2013; Skinner & Belmont, 1993). Providing structure and guidance has been found to enhance important outcomes including students' behavioral engagement (e.g., effort and persistence), positive emotions, and use of effective learning strategies (e.g., Jang et al., 2010; Mouratidis et al., 2013; Skinner & Belmont, 1993). Teachers can foster students' interest and enjoyment in mathematics and other subjects by using clear rules and monitoring strategies during mathematics lessons (Kunter, Baumert, & Köller, 2007). High quality instruction involves providing clarity and structure and employing non-authoritarian teaching styles and has been shown to reduce mathematics anxiety, which shares features with emotional cost (Frenzel, Pekrun, & Goetz, 2007; Middleton & Spanias, 1999). A high level of structuredness and extensive support have been found to have a positive association with students' self-concept in mathematics class at secondary school (Lazarides & Ittel, 2012). Interestingly, at secondary school, structure is most effective in combination with the establishment of positive relationships (e.g., social support) or autonomy support (Jang et al., 2010; Lazarides & Ittel, 2012). Teaching approaches relying on minimal guidance have, in turn, been found to lead to higher emotional costs (e.g., frustration) and fewer learning gains, especially in STEM subjects (Kirschner, Sweller, & Clark, 2006).

Sometimes studied as a part of teaching that is autonomy-supportive (Urdan & Schoenfelder, 2006) or that focuses on giving structure (Jang et al., 2010), **using appropriate feedback strategies** also has been considered an essential element of motivational instruction (e.g., Wentzel & Brophy, 2014). Indeed, giving students informative, constructive feedback and helping them find the most appropriate and effective learning strategies have been found to improve students' intrinsic motivation, academic competence beliefs (e.g., self-efficacy, self-esteem, self-concept), and behavioral engagement in the classroom, to reduce students' emotional costs, and to increase long-term engagement in learning (for reviews, see e.g., Hattie & Timperley, 2007; Henderlong & Lepper, 2002; Wentzel & Brophy, 2014). Feedback that includes intraindividual comparison referring to prior success (i.e., verbal persuasion) or positive social comparison referring to the success of equally able peers (i.e., modeling) is considered a powerful tool to reinforce students' self-efficacy beliefs after failure on a given task (e.g., Schunk & Zimmerman, 2007; Urdan & Schoenfelder, 2006). Especially in mathematics and science, feedback has been found to be particularly effective when it refers more to effort than to ability (e.g., Dweck, 2008; Wentzel & Brophy, 2014).

In summary, empirical research in education has revealed a number of instructional methods with beneficial effects on students' motivation and engagement. However, numerous studies based on SDT (Ryan & Deci, 2000) and the assumption that fulfilling students' basic psychological needs for autonomy, social relatedness, and competence in the classroom has the power to turn extrinsic into intrinsic motivation (Ryan & La Guardia, 2000) have focused on the effects of teaching on students' interest or intrinsic motivation, whereas other value beliefs (attainment, utility, cost) have been ignored. In addition, research on the motivating effect of relevance-enhancing teaching strategies has been comparatively scarce (though sometimes assessed as part of overall autonomy-supportive teaching; e.g., Skinner & Belmont, 1993). Interestingly, researchers and practitioners in education recently have called for mathematics and science instruction in which relevance is emphasized (e.g., Davis & McPartland, 2012; Osborne & Dillon, 2008), primarily as a result of the continued need for qualified workers in STEM fields. Thus, the assumed motivational potential of relevance-oriented teaching strategies is explored in the following paragraph.

### 2.1.2 The potential of relevance-oriented teaching

> **The experience of meaningfulness of mathematics [is neglected] (...). Few now deny that school mathematics as experienced by most students is compartmentalized into meaningless pieces that are isolated from one another and from the students' wider world. Symbols are manipulated without regard to the meanings that might be carried, either by referents of the symbols or by actions on them. Theorems are 'proved' without the slightest attempt (...) to justify the need for proof. This experienced meaninglessness of school mathematics devastates the motivation to learn or use mathematics and is entirely incompatible with a view of mathematics as a tool for personal insight and problem solving.**
>
> **(Kaput, 1989, pp. 99-100)**

Helping students see the relevance of learning topics covered in the curriculum is a central task in education; however, this task seems to be challenging. In mathematics lessons in particular there is a certain risk that students perceive mathematics as a list of rules and formulas that must be memorized but are of little importance to students' personal lives (e.g., Kaput, 1989; Osborne & Dillon, 2008). Consequently, students' motivation to learn and apply mathematical rules and formulas may suffer. Indeed, reviewing research on students' perceptions of the usefulness of learning mathematics (see **1.3.4**) it seems that even almost thirty years after Kaput's (1989) criticism about mathematics instruction there still is an important number of students for whom many mathematics topics are "meaningless pieces". Qualitative research underlines this assumption: For several mathematical domains, secondary school students find it difficult to give examples of when their mathematical knowledge is useful in real-life situations (Harackiewicz, Hulleman, Rozek, Katz-Wise, & Hyde, 2010).

It is mathematics teachers who, as important socializers of their students, bear a remarkable responsibility to help students connect the learning material in mathematics to their everyday lives and thereby to support learners' motivation. EVT (Eccles et al., 1983) and other theories (e.g., stage-environment fit theory, Eccles, Midgley et al., 1993; control-value theory of achievement emotions, Pekrun, 2006) assume that through their teaching behavior and own values, mathematics teachers may become mediators of students' general value development and, in turn, affect students' learning attitudes and behavior. According to Pekrun (2006), teachers can foster the positive value of engaging in academic activities in students through direct verbal messages as much as through their instructional behavior and choice of learning assignments. For instance, teachers who intend to convey to their students the relevance of mathematics topics covered in class may talk about everyday situations in which students may be able to apply their knowledge of mathematics or they may choose activities with a great personal importance in their students' daily lives. Therefore, there is an inseparable link between the teacher's values and the characteristics of the learning activities.

For decades, researchers and educationists have made strong claims to "teach mathematics so as to be useful" (title of an essay by Freudenthal, 1968), have given recommendations about "teaching mathematics meaningfully" (title of a book by Allsopp, Kyger, & Lovin, 2007) and have defined the support of relevance and meaningfulness as a core motivational principle in the classroom (Davis & McPartland, 2012; Kaput, 1989; Keller, 1983; Osborne & Dillon, 2008; Pianta et al., 2012; Turner, Warzon, & Christensen, 2011; Wentzel & Brophy, 2014). From the perspective of SDT, the provision of a meaningful "rationale" for learning also has been assumed to foster students' motivation when it is part of an autonomy-supportive teaching style (e.g., Skinner & Belmont, 1993). Indeed, results of laboratory studies indicate that giving students rationales for learning topics increases undergraduates' engagement in academic activities they find uninteresting and promotes subsequent deep-level learning (Deci, Eghrari, Patrick, & Leone, 1994; Jang, 2008; Reeve, Jang, Hardre, & Omura, 2002). In one of the rare studies of "fostering relevance" as a separate teaching strategy in secondary school classrooms, explaining relevance was found to be more predictive of positive affect and behavioral and cognitive engagement than other autonomy-supportive strategies (Assor et al., 2002). Wang (2012) demonstrated a positive influence of teaching for meaning in mathematics in Grade 7 on a composite measure of students' mathematics-related values in Grade 10. Furthermore, Lazarides and Rubach (2017) found an indirect positive link between teaching for meaning and students' intrinsic motivation to learn mathematics in Grade 10.

However, several questions remain open with regard to the motivational potential of relevance-oriented teaching in mathematics class. First, previous studies have failed to include all value dimensions as outcomes, as either a composite value measure was used (e.g., Wang, 2012) or focus was on only one value dimension such as students' interest (e.g., Lazarides & Rubach, 2017). Moreover, despite its educational importance, the cost associated with learning mathematics often has been neglected in research on teachers' impact on students' value beliefs (cf., Wigfield & Cambria, 2010). Second, relevance-oriented teaching has been investigated only on a relatively general level without referring to specific teaching strategies conveying the relevance of learning mathematics topics. It would be important for teachers to know which kinds of strategies are most effective in enhancing students' value beliefs. Lastly, focusing mainly on the teacher, the social complexity of the classroom has not been considered in prior studies. Peers also may shape students' values in the classroom, especially during the teenage years (e.g., Eccles, Midgley et al., 1993; Kindermann, 2007). Accordingly, to evaluate the effectiveness of relevance-oriented teaching strategies, students' perception of classmates' value-related behavior also needs to be taken into account (see **2.1.3**).

What do instructional strategies to convey the relevance of mathematics topics entail? The basic idea of establishing relevance in teaching mathematics is that the teacher helps students recognize the relationship between the tasks done in mathematics class and their personal life (e.g., Wentzel & Brophy, 2014). In addition to stressing the practical applicability of mathematical rules and formulas in students' everyday lives (e.g., Rakoczy, Klieme, & Pauli, 2008), another relevance-oriented strategy involves using examples from students' everyday lives when introducing new topics in mathematics class (e.g., Freudenthal, 1968). Furthermore, results of qualitative studies (e.g., Michelsen & Sriraman, 2009) indicate that highlighting connections between mathematics and other school subjects can have positive effects on students' perception of the relevance of learning mathematics topics and their interest in mathematics, as mathematics skills are shown to be relevant in other academic domains. In summary, stressing the practical applicability of mathematical rules and formulas, introducing new mathematics topics with everyday examples, and demonstrating links to other academic domains seem to be appropriate instructional strategies to convey the relevance of mathematics topics and to thus support students' mathematics-related value beliefs.

### 2.1.3 Peers' influence on students' motivation to learn mathematics

Learning at school is a social endeavor, and the classroom is a socially complex environment in which the teacher and his or her teaching strategies as well as peers may play an important role in students' development of values. Especially during puberty, the focus of young adolescents' social life shifts away from relationships with adults toward relationships with peers: Popularity among classmates becomes increasingly important (Simmons & Blyth, 1987). Researchers have argued that this heightened peer orientation needs to be taken into account when investigating the effects of the classroom context on motivational variables (cf., stage-environment fit theory by Eccles, Midgley et al., 1993). More particularly, peers have been found to influence general academic engagement in various ways (for reviews, see Fredricks et al., 2004; Juvonen et al., 2012). Classmates are believed to shape the classroom context by bringing in norms and values. Through interaction in the classroom, students experience their classmates' attitudes on a daily basis (Kindermann, 2007; Ryan, 2000). Classmates may have a strong influence on each other's values especially in school systems where students stay together with the same classmates for several subjects and from grade to grade (cf., Frenzel, et al., 2010).

Although it is obvious that peers are important to students at secondary school, the focus of empirical research on students' motivation and engagement in the classroom context has long been the impact of the teacher (cf., Fredricks et al., 2004; Juvonen et al., 2012). In more recent empirical studies, however, the influence of peers on students' motivation, interest, and behavior has been acknowledged. For instance, Goodenow and Grady (1993) found positive associations between students' perceptions of the value their friends attribute to schoolwork and students' own academic competence beliefs and value beliefs (composite measure of interest and importance). Nelson and DeBacker (2008) found positive associations between classmates' achievement-related norms and the value their best friends placed on learning with students' mastery goals and social norms in science class. However, cross-sectional research does not allow for causal inference (cf., Shadish et al., 2002).

Results of longitudinal studies indicate that throughout secondary school, students' level of academic motivation becomes more and more similar to those of their peers (Kindermann, 2007; Kindermann, McCollam, & Gibson, 1996). In addition, Ryan (2001) found the peer group value climate to be related to changes in intrinsic value of schoolwork and academic achievement for seventh-graders over the school year, but it did not predict changes in students' perception of the importance and usefulness of their schoolwork. Because Ryan (2001) analyzed only peers' mutual influence on values related to schoolwork in general, it is not known whether her findings would be the same for mathematics in particular. Furthermore, Frenzel et al. (2010) investigated interest development in mathematics from Grade 5 through Grade 9. They observed that at both the individual and class levels students' perception of the value their classmates attributed to learning mathematics was associated positively with changes in students' interest in mathematics. However, they did not investigate the influence of peers on changes in other value beliefs. In summary, ostensibly no longitudinal studies have been conducted of the influence of teachers' and peers' value-related behaviors on the simultaneous development of students' separate value beliefs (intrinsic, attainment, utility, and cost).

## 2.2 Scientific interventions aimed at fostering students' motivation

Longitudinal observational data collected within schools are an important source to learn more about students' motivational development in genuine classroom environments and thus provide empirical evidence for psychological hypotheses and theories. However, experimental studies are valuable for identifying causal relationships, that is, to attribute an effect to a cause (Shadish et al., 2002). Through intervention studies it is possible to learn more about what happens, for instance, to students' effort after their perception of the relevance of learning about a mathematics topic has been manipulated intentionally. In addition, when students engage in intervention activities requiring the production of texts, these texts can be analyzed to find causal explanations clarifying the processes through which the intervention produced an effect (e.g, Nelson, Cordray, Hulleman, Darrow, & Sommer, 2012; Shadish et al., 2002). Therefore, experimental studies are essential to advancing psychological theorizing and to finding ways to improve instructional practices. Such use-inspired basic research providing information for both educational theory and practice has been claimed to be key to innovative motivational research (Pintrich, 2003).

Many scholars researching motivation have developed, implemented, and evaluated educational interventions designed to boost students' motivation such as relevance interventions (for recent reviews, see Hulleman et al., 2016; Karabenick & Urdan, 2014; Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016). In the following chapter an overview is given of intervention approaches tested in school classrooms which aim to improve students' competence beliefs and value beliefs in mathematics or science, with a particular focus on the potential of relevance interventions. Subsequently, a closer look is taken at the psychological processes at work in relevance interventions—and the potential of relevance interventions to contribute to both educational practice and psychological theorizing.

### 2.2.1 Intervening on students' motivation to learn mathematics: the potential of relevance interventions

In numerous experimental studies interventions have been conducted in the classroom on students' **competence beliefs** in mathematics or science but have produced mixed results. Some approaches such as supporting students' experience of mastery through achieving self-set goals (Schunk, 1985), having students explain the topic or concept to themselves (Crippen & Earl, 2007), and encouraging students through verbal persuasion (e.g., providing attributional feedback, e.g., Craven, Marsh, & Debus, 1991; Schunk, 1983) have improved students' self-efficacy, self-concept, and test-based achievement in mathematics (for meta-analyses of self-concept/self-esteem interventions, see Haney & Durlak, 1998; O'Mara, Marsh, Craven, & Debus, 2006; for reviews of self-efficacy interventions, see Schunk & Ertmer, 1999; Schunk & Mullen, 2012). However, numerous interventions on students' competence beliefs, especially in secondary school mathematics or science, did not result in any significant improvement in students' self-efficacy or self-concept (e.g., Brewer & Becker, 2010; Hodges, 2008; Isiksal & Askar, 2005; Ramdass & Zimmerman, 2008; also see review by Rosenzweig & Wigfield, 2016). After all, students' competence beliefs develop over years and are relatively stable, especially students' academic self-concepts (e.g., Schunk & Pajares, 2009). Accordingly, intervention programs targeting students' self-concept typically last several months (cf., meta-analysis by O'Mara et al., 2006). In addition, sustained improvement in students' self-concept through scientific interventions is more likely when there is an actual improvement in students' achievement (cf., O'Mara et al., 2006). In summary, the malleability of competence beliefs—in particular students' self-concept—and the success of respective scientific interventions depend on a complex interplay of various factors.

Concerning students' **value beliefs**, few interventions have targeted students' intrinsic and attainment value beliefs or cost. Although various classroom-based motivational interventions have measured interest or intrinsic value as an outcome (e.g., Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009; see also Rosenzweig & Wigfield, 2016), few experiments intervened explicitly on interest or intrinsic value. For instance, Walkington (2013) conceptualized an intervention which was based directly on students' personal interests: The author manipulated the students' learning environment in mathematics class at secondary school by matching the content of algebra problems to students' preferred leisure activities. Compared to students in the control group, who did not receive personalized mathematics problems, students in the interest condition were faster at solving algebra problems, performed better, and processed the learning contents more deeply—even when the intervention was removed. Similarly, cost rarely has been a direct target of classroom-based interventions—and if so, the focus has been mainly on emotional cost such as mathematics anxiety. For instance, letting students write about their feelings before an upcoming examination (Ramirez & Beilock, 2011) or telling students that arousal actually has a positive effect on test achievement (Jamieson, Mendes, Blackstock, & Schmader, 2010) have been shown to be effective approaches to decreasing students' mathematics anxiety and to improving their performance in mathematics.

In previous value intervention studies focus was mainly the topic of personal relevance and, thus, students' utility value beliefs were investigated. In fact, students' perception of relevance seems to be more amenable to scientific interventions than all other value components for several reasons. Whether a student enjoys engaging in mathematics-related tasks or not seems to depend very much on individual preferences and characteristics, which are potentially difficult to trigger. Elaborating on more rational reasons a subject is relevant for a student's life, however, may be a feasible way to foster perceptions of relevance and meaningfulness. Compared to intrinsic value and attainment value, utility value is more extrinsic in nature (Eccles & Wigfield, 2002) and can be influenced rather easily from the outside, which makes it a promising target for classroom interventions (e.g., Trautwein et al., 2013). However, utility value is related to personal goals and therefore also has an intrinsic component (Eccles, 2005), which is why intervening on relevance also has the potential to foster more intrinsic aspects of students' motivation. More precisely, several authors have argued that initial stimulation of more extrinsic motivation should lead to enhanced engagement and thereby promote students' intrinsic motivation and interest development (Hidi & Harackiewicz, 2000; Hidi & Renninger, 2006). Further, incorporating beliefs about the relevance of a task into one's concept of self might have a positive effect on students' confidence in their ability to do well on a given task and thus support students' competence beliefs (Husman & Lens, 1999; Schunk & Mullen, 2012).

In previous research these assumed benefits of intervening on students' perceptions of relevance have started to be explored while taking into consideration Walton's (2014) three principles of "wise" psychological interventions. Walton (2014) suggests that educational interventions are wise when they are, first, psychologically precise, that is, based on formalized strategies derived from social-psychological theory and rooted in laboratory research. Accordingly, researchers have formulated explanations framed in EVT (Eccles et al., 1983) for how various intervention approaches may affect students' utility value beliefs. They also have formulated theories based on results from laboratory research about how different intervention approaches may affect students' utility value beliefs (see **2.2.2**). In addition, researchers recently have started to investigate more closely the specific mechanisms through which changes in perceptions of relevance are mediated (see **2.2.4**). Second, psychological interventions need to be tested for context dependency through field experiments. Researchers have begun to evaluate the effectiveness of relevance interventions in genuine classroom contexts (see **2.2.3**). Third, psychological interventions should target recursive processes to cause lasting change. Indeed, by initiating (repeated) reflections on the relevance of learning tasks, relevance interventions are assumed to facilitate later positive personal associations with the subject matter and subsequently lead to other positive academic outcomes. Although there is a solid research base on relevance interventions, little is known about the breadth and sustainability of the effects, and further questions remain unanswered (see **2.2.5**).

### 2.2.2 Relevance interventions in the laboratory

Research on relevance interventions started in the laboratory with a range of experiments in which college students were asked to study a new multiplication technique (for a review, see Durik, Hulleman, & Harackiewicz, 2015). The first approach to conveying the relevance of mathematics was rather straightforward: Students were provided with information about the usefulness of learning the new multiplication technique for everyday situations such as banking, tipping at restaurants, and calculating discounts at retail stores (Durik & Harackiewicz, 2007, study 2). Compared to students in the control condition, who had not read about the usefulness of the technique, students in the relevance condition reported greater interest and higher attainment value (competence valuation) concerning the technique. However, the intervention had a positive effect on students' interest, attainment value, and engagement in a subsequent problem-solving task only when the students' initial interest in mathematics was already high. In contrast, for students with initially little interest in mathematics, no effects were found on value beliefs and negative effects were found on competence beliefs.

Using these findings, Hulleman et al. (2010, study 1) designed another laboratory intervention in which college students were taught the same multiplication technique. Students in the experimental condition—instead of being given information about the relevance of the technique—were asked to generate examples of how this multiplication technique related to their personal lives. Compared to students in the control condition, who wrote an unrelated essay, this relevance intervention enhanced students' utility beliefs of, and interest in, the technique. Interestingly, the intervention was found to be particularly effective for students with negative competence beliefs in connection with the technique.

In a series of lab studies, Canning and Harackiewicz (2015) first replicated results from previous laboratory intervention studies (Durik & Harackiewicz, 2007; Hulleman et al., 2010) and then compared the effectiveness of the interventions when relevance was explained to students and when students identified relevance themselves. They found that when relevance was explained to students, the interventions reduced their interest and performance if students had negative competence beliefs, and when students identified relevance on their own, the interventions had positive effects. Moreover, the authors found synergistic effects of directly communicating relevance and having students identify relevance: Presenting information about the usefulness of the task first and then asking students to generate personal arguments about its usefulness was found to be an effective way to boost students' utility value beliefs, task interest, and performance independently of students' competence beliefs. Lastly, the authors found that directly-communicated everyday examples of the relevance of the task were more effective than school- or career-related examples.

### 2.2.3 Classroom-based relevance interventions

Hulleman and Harackiewicz (2009) conducted a randomized field experiment with ninth-grade students during science class at school which was based on findings from the above outlined laboratory studies. Students were randomly assigned to either the relevance condition or the control condition. All participating students were asked to write eight essays throughout one semester: students in the relevance condition wrote about the meaning of the course material to their lives; students in the control condition wrote summaries of the course material. Students' interest and grades in the science class were promoted through the intervention when students had negative competence beliefs, whereas students with positive competence beliefs did not differ in their outcomes from highly confident students in the control group, a replication of findings from their previous laboratory study (Hulleman et al., 2010).

Similar results have been obtained in further studies with undergraduate students of psychology (Hulleman et al., 2010, study 2; Hulleman, Kosovich, Barron, & Daniel, 2017, study 2) and biology (Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016). All studies had a similar experimental design: In class, students were assigned to either a relevance condition (e.g., writing about the personal relevance of the topic) or to a control condition (e.g., summarizing the topic) and wrote up to three essays during the semester. The relevance intervention conditions fostered the interest (Hulleman et al., 2010), competence beliefs (measured through success expectancy; Hulleman et al., 2017), and grades (Harackiewicz et al., 2016; Hulleman et al., 2017) of all students in the course. In addition, students with little confidence and/or low performers profited from the relevance interventions with regard to course-related interest (Hulleman et al., 2010; 2017), utility value, competence beliefs (Hulleman et al., 2017), and grades (Harackiewicz et al., 2016; Hulleman et al., 2017). Furthermore, Harackiewicz et al. (2016) found the relevance interventions in particular enhanced the achievement of first-generation underrepresented minority students, thus closing social-class related achievement gaps in college biology courses.

In addition to the relevance studies based on essay-writing conducted by Harackiewicz and Hulleman, further approaches to conveying the relevance of mathematics topics and science topics to students in the classroom have been developed and tested. For instance, Woolley et al. (2013) conducted a school-level intervention study of middle-school students who were surveyed during 10 mathematics lessons per year from Grade 6 to Grade 8. During these lessons teachers provided students with a variety of standardized career-relevant examples and problems during instruction while adhering to the standard curriculum for mathematics. Relevance-oriented mathematics instruction was found to promote students' final grades in mathematics at the end of Grades 7 and 8—an effect which was not found at the end of Grade 6.

Finally, with a more complex intervention design, Acee and Weinstein (2010) targeted college students' utility value as well as their interest and attainment value in statistics, using several short sessions of reading passages and subsequent writing activities. In the reading passages, students read, for instance, about the personal relevance of statistical knowledge, about the importance of a positive attitude towards statistics, or how learning statistics can be both challenging and interesting. In the writing activities, students were guided, for instance, to relate statistics knowledge and skills to their future education and career, to construct a positive attitude toward statistics activities, or to find solutions how to overcome obstacles in learning statistics. Positive effects were found on students' immediate and 2-week delayed measure of students' task value (combined scale of students' intrinsic, attainment, and utility value) and utility of learning the course material for students' personal future, but no effects were found on students' self-efficacy. Furthermore, students' continued interest measured through choice-behavior and course performance was boosted through the intervention.

### 2.2.4 Toward causal explanations: How do relevance interventions work?

The success of the aforementioned relevance interventions to improve student learning outcomes in the laboratory and in the classroom is compelling: Enhancing students' utility value beliefs boosted students' value beliefs (interest, attainment, utility), competence beliefs (success expectancy), behavior (task engagement), and—partially—performance (task performance or grades). This body of experimental research showing that the explicit manipulation of students' utility value beliefs has an effect on important student outcomes constitutes a valuable basis upon which causal relationships can be inferred and described (Shadish et al., 2002). However, relevance interventions, especially classroom-based ones, have not always produced consistent effects. For instance, students' performance (scores on examinations or grades) was boosted directly through the relevance interventions by Woolley et al. (2013) and Harackiewicz et al. (2016). In other studies effects on students' performance were reported for students with little confidence or underperformers only (Hulleman & Harackiewicz, 2009; Hulleman et al., 2017) or no effects on performance were reported at all (Hulleman et al., 2010). Sometimes, main effects of interventions on students' interest were found (e.g., Acee & Weinstein, 2010; Hulleman et al., 2010); sometimes effects were moderated (Hulleman & Harackiewicz, 2009; Hulleman et al., 2017). As only a selection of outcome measures and only effects within a subsample of intervention groups (Harackiewicz et al., 2016) were reported in the publications, there might be further inconsistencies in the research on the effectiveness of relevance interventions which are concealed due to publication bias (cf., Schmiedek, 2016). Further, relevance interventions sometimes do not produce any effect at all (e.g., Husman, Nelson, & Cheng, 2017; Karabenick, Albrecht, & Rausch, 2017).

What lies behind these differences in effects? There were differences in the design of the interventions (settings, populations, intervention tasks, durations, and evaluation designs) and in the measurement of the outcomes (e.g., grades vs. standardized measures of performance) which have to be considered in the interpretation of the results. However, to be able to explain how relevance interventions affect students' motivation, behavior, and/or performance, mediating processes have to be identified. In other words, studies providing causal explanations for the effectiveness of relevance interventions are needed (cf., Shadish et al., 2002). Furthermore, knowing the psychological processes that do or do not lead to an effect on the outcome is important to be able to adapt these interventions to different contexts and populations (cf., Walton, 2014). To do so, it is important to analyze which parts of the interventions have actually been put into practice—a matter of intervention fidelity (Dane & Schneider, 1998).

*Intervention fidelity* refers to whether the actual intervention implementation corresponds with the theoretically planned intervention program (Dane & Schneider, 1998). Especially in genuine educational settings as in school classrooms, it is difficult to ensure that all theoretically planned parts of an intervention have been implemented in practice as intended (O'Donnell, 2008).Variation in intervention fidelity may be attributable to the implementers (e.g., teachers) and, for instance, their adherence to the intervention manual, as well as to the participants (e.g., students) and the extent to which they are engaged in the intervention activities (i.e., participant responsiveness; Dane & Schneider, 1998). If teachers or researchers implement only a part of the theoretical intervention program, the effects likely are due to the implemented parts rather than the entire intervention program. Similarly, if students do not respond to all theoretically important aspects of an intervention activity, it is those aspects students actually respond to that are likely to underlie the observed intervention effects. Investigating intervention fidelity is essential to understanding the processes of determining why an intervention is effective or not, or why it works for certain subgroups of students only (Murrah, Kosovich, & Hulleman, 2017).

One aspect of intervention fidelity that is particularly central to relevance interventions is *students' responsiveness* to the intervention, which is the degree to which students are involved in the intervention activities (Dane & Schneider, 1998). As in many social-psychological interventions, a personal writing activity is at the heart of relevance interventions in the classroom (Yeager & Walton, 2011). In all of the aforementioned classroom-based relevance interventions except for the study by Woolley et al. (2013) students were asked to complete writing tasks about the relevance of science-related topics. The process of writing about relevance is expected to trigger a change in students' motivational beliefs. However, how do these writing activities have to be completed to uncover their full potential on students? What are the key indicators students have to respond to in their essays to trigger a change in their perceptions of relevance? How well do students respond to these indicators? What are the characteristics of students who respond particularly well to the writing tasks—or particularly badly? And how does the degree of students' responsiveness relate to the strength of the intervention effects? Answering these questions is crucial to be able to optimize intervention designs and thus pave the way for relevance interventions to enter classrooms on a large scale (Cohen & Loewenberg Ball, 2007).

Preliminary insights into the indicators that are central to making interventions successful have been gained through mediation models. Hulleman et al. (2010) investigated whether the effect of the relevance intervention on students' interest and achievement could be explained through students' utility value beliefs, the target outcome of relevance interventions. Indeed, perceived usefulness mediated the effects on interest so that the effect of the intervention was no longer significant in either the laboratory study or the college classroom study. In addition, students' utility value mediated the effect of the relevance intervention on students' performance in the classroom study but not in the laboratory study. This mediation, however, could not be replicated with the data of later relevance interventions (Hulleman et al., 2017). In addition, it remains unclear how the change in students' utility value beliefs was triggered through the intervention.

In a relevance experiment conducted in undergraduate biology class Harackiewicz et al. (2016) analyzed students' responsiveness to the writing activity by counting the length of students' relevance essays and coding the words indicative of social processes, personal connections, and cognitive involvement. Mediation models were reported for essay length only, showing that the number of words in students' relevance essays mediated the intervention effects on students' course achievement. However, first generation underrepresented minority students, who profited most from the intervention, used more words indicative of social processes and cognitive involvement, which caused the authors to argue that these two aspects of students' responsiveness contributed to the intervention effects. Furthermore, the authors found that students with high initial achievement wrote the longest essays and made most personal connections.

Using a descriptive analytical approach to determine the relative intervention strength achieved, Hulleman and Cordray (2009) investigated why the effects of the laboratory relevance intervention (Hulleman et al., 2010, study 1) on students' utility value beliefs could not be replicated in the relevance intervention in secondary school science class (Hulleman & Harackiewicz, 2009). Responsiveness was assessed by the extent to which the students connected the mathematics activity or science topic to their personal lives in their essays. Analyses showed that the number and quality of personal connections to the learning material made by students in the classroom was lower than those made by students in the laboratory. Similarly, in a later experiment, Hulleman et al. (2017) found that students in the relevance interventions made more personal connections in their essays than students in the control condition, leading to the authors' conclusion that the frequency of connections was at the heart of the success of relevance interventions. To learn more about the contribution of personal connections to the intervention effect, the authors attempted to manipulate the number of connections students made with the learning material experimentally. The experimental manipulation, however, was not successful.

### 2.2.5 Driving relevance intervention research forward

In summary, relevance interventions have been shown to be an effective tool to enhance students' value beliefs and—partially—improve grades when conducted in the classroom, especially when students have been asked to write about how science-related subjects are relevant to their personal lives. In addition, attempts have been made to investigate the processes underlying the intervention effects, indicating that factors such as quality and quantity of personal connections could be important for relevance interventions to work. However, further research is needed to advance research on relevance interventions so that findings can be used to improve educational practice and psychological theories and eventually enter educational practice.

First, successful relevance intervention approaches need to be replicated in different settings with **different populations**. To scale up psychological experiments in education, Walton (2014) recommended they be replicated within and adapted to diverse educational settings and populations. In fact, it is striking that all of the aforementioned studies were conducted in the United States and have not yet been replicated with students in other nations. Replication is of paramount scientific importance so as to avoid a distorted picture of intervention effectiveness (cf., Schmiedek, 2016). Although the focus of a range of laboratory experiments has been mathematics-related relevance, almost none of the classroom-based relevance interventions have been conducted in the subject of mathematics (see Woolley et al., 2013, for an exception). Successful interventions, for example those in which students write about the personal relevance of a topic in a text, still have to be tested with a non-American sample and in the subject of mathematics.

Second, **new intervention approaches** to conveying relevance need to be developed. Integrating the provision and the self-generation of relevance information into one intervention approach has been shown to be successful in a laboratory setting (Canning & Harackiewicz, 2015). Acee and Weinstein (2010) included a reading passage on the relevance of learning statistics and a subsequent writing activity to have students generate their personal ideas about relevance in a comprehensive value intervention. However, such a combined approach has not yet been investigated in a classroom setting for targeted relevance interventions. In addition, further approaches to conveying the personal relevance of topics need to be tested and compared with conventional approaches to find out which relevance interventions have the strongest and most sustained effects. In fact, a relevance intervention targeting secondary school students' parents has shown that mailing information about the usefulness of STEM courses (through brochures on a Web site) to students' parents caused secondary school students to continue taking mathematics and science courses at secondary school for a longer time than students whose parents had not received the respective information (Harackiewicz, Rozek, Hulleman, & Hyde, 2012). Just as the parents communicated their beliefs about the relevance of mathematics or science to their kids, peers or young adults could act as role models for secondary school students and convey the usefulness of mathematical knowledge by describing situations in which they needed mathematical skills.

Third, relevance interventions need to be **adapted to students' genuine classroom environment**. Previous relevance intervention studies have been conducted mainly at the individual level, assigning students within classes to different conditions. Such an experimental design has the advantage that variation across classes (e.g., resulting from the teaching style of the teacher) can be kept constant and that at the student level, smaller sample sizes are sufficient to reach a reasonable statistical power. However, within-class randomization also has drawbacks: First, the risk of diffusion effects between students in different conditions within a class is high (Craven, Marsh, Debus, & Jayasinghe, 2001); second, the setting does not correspond with students' genuine classroom setting where students within classes are typically allowed to discuss their common learning experiences. Conducting relevance interventions at the class level opens new possibilities such as initiating class discussions about the relevance of mathematics or science.

Fourth, **short- and long-term effects on neglected outcomes** need to be investigated. Wise interventions target recurring psychological processes in order to bring about lasting change (Walton, 2014). It is unclear how students' motivation and achievement develops once the intervention is removed, as there have been few follow-up studies. Until now, classroom-based relevance interventions have targeted mostly task-related relevance (i.e., writing about the usefulness of a specific topic), but an effective method to have students make personal connections on their own on a regular basis (i.e., for further course topics) has not yet been developed (cf., Hulleman et al., 2017). Targeting students' general, domain-specific relevance beliefs (instead of task-specific beliefs) could be another way to bring about lasting change. If beliefs about the present and future relevance of a domain are enhanced, this change should affect students' motivation over and above the topic actually addressed in class. However, to be able to make claims about the sustainability of intervention effects, follow-up studies are needed in which evaluation is made of the effects not only on students' achievement (e.g., Hulleman et al., 2017; Woolley et al., 2013), but on a range of important outcomes. Behavioral measures (e.g., effort), motivational measures (e.g., competence beliefs), and test-based achievement (instead of grades, which measure more than mere achievement, e.g., McMillan, 2001) have been understudied in prior relevance intervention research. Furthermore, investigation into the effects of relevance interventions on both subject-specific competence beliefs and task-specific competence beliefs is needed to understand better the nature of the relationship between competence beliefs and value beliefs (see **1.3.2**).

Finally, **in-depth fidelity analyses** need to be conducted to unravel the processes mediating the effects of relevance interventions. To be able to make precise adaptations of relevance interventions to different settings and populations, and eventually to pave the way for relevance interventions to enter educational practice, more knowledge is needed about how relevance interventions work, why they sometimes do not work, or why they work for certain students only (e.g., Cohen & Loewenberg Ball, 2007; Murrah et al., 2017; Walton, 2014; Yeager & Walton, 2011). In previous intervention research assumptions were made about the importance of different elements of students' responsiveness (e.g., personal connections, cognitive involvement) to the writing activities (e.g., Harackiewicz et al., 2016; Hulleman & Cordray, 2009). However, no comprehensive analysis has been conducted of students' responsiveness to the interventions, including the assessment of several indicators of responsiveness, their predictability from students' individual characteristics, and both descriptive and causal approaches to analyzing their contribution to the intervention effects. More precisely, identifying individual characteristics that predict students' responsiveness is necessary to determine possible differential intervention effects and to find ways of adapting the intervention material to the needs of less responsive students, thereby potentially increasing the intervention effects (Nelson et al., 2012). Previous research on the importance of students' responsiveness for the effectiveness of relevance interventions has been descriptive (e.g., Hulleman & Cordray, 2009), and attempts to manipulate experimentally single indicators of responsiveness (e.g., connection frequency) have failed (Hulleman et al., 2017). Including descriptive information on single indicators of responsiveness and conducting in-depth analyses of causal effects based on students' responsiveness to the writing activities is needed to understand better how single indicators and different degrees of students' responsiveness actually matter for the intervention effects.

# 3

# Aims and research questions

### 3.1 Background and aims of the empirical studies

Fundamental knowledge of mathematics is useful in everyday life; it is needed to participate actively in society and to qualify for many professions (Joint Economic Committee, 2014; Klein & Schimmack, 1907; OECD, 2016b). Although the personal relevance of mathematics is uncontroversial among educationists (Klein & Schimmack, 1907), students still have difficulty seeing the everyday usefulness of mathematics (e.g., Reiss et al., 2016). In this dissertation investigation is made into the impact of emphasizing the relevance of mathematics on students' motivation, behavior, and achievement in mathematics. The relevance of mathematics is conveyed through either teachers and peers or scientific interventions during mathematics class.

Framed within Eccles et al.'s (1983) EVT of achievement motivation, this dissertation builds upon a large body of research underlining the importance of students' competence beliefs and value beliefs, including perceived relevance, for their academic behavior and achievement. Although the focus of research on students' motivation has been mainly the effects of students' competence beliefs and value beliefs on their academic outcomes (cf., e.g., Wigfield et al., 2009), empirical investigations into how students' motivational beliefs are affected by their socializers' beliefs and behaviors have been comparatively scarce. In particular, no studies seem to have been conducted of the impact of teachers and classmates and their conveying of relevance and value on all components of students' value beliefs. Results of scientific interventions designed to foster students' value beliefs in the classroom have indicated the compelling potential of relevance interventions; however, few approaches have been researched, few outcomes have been explored, and the processes through which these interventions work (or do not work) rarely have been investigated in depth. More research is needed to determine which relevance interventions are the most effective and why they work (or not), and how relevance interventions can be adapted for different settings and samples.

This dissertation expands on previous correlational and experimental research framed within EVT as the importance of conveying the relevance of topics addressed in mathematics class is explored in several new ways. First, investigation is made into teachers' instructional strategies and students' perceptions of their classmates' value-related behavior in order to determine how students' value beliefs are affected by relevance-oriented teaching in everyday classroom practices. Second, knowledge of the potential of scientific interventions targeting students' perception of the relevance of mathematics is extended by a) developing new relevance intervention approaches, b) implementing them with a new sample at the classroom level, and c) evaluating and comparing their effectiveness on a broad range of previously neglected outcomes including students' motivational beliefs, effort, and achievement. Third, new knowledge of the processes underlying the effects of classroom-based relevance interventions is generated by taking a closer look at students' responsiveness to the intervention activities, investigating the characteristics of responsive and nonresponsive students, and analyzing the role of students' responsiveness for the effectiveness of relevance interventions. The current research eventually will integrate the newly gained knowledge about the potential of relevance-oriented teaching and intervening on relevance to suggest effective practice-oriented ways to boost students' motivation in mathematics class.

This dissertation contributes to the field of use-inspired basic research (Stokes, 1997) as investigation is made into the importance of students' perception of the relevance of mathematics topics addressed in class resulting either from everyday teaching practices or from interventions developed by researchers. The results of this dissertation are useful for educational practice because knowledge derived from EVT (Eccles et al., 1983) is applied to genuine classroom contexts to investigate the importance of the theory-driven element of "relevance perception" and to foster students' motivation, behavior, and achievement. The results contribute to the understanding of the relationship among various social-cognitive factors related to students' motivation, for instance, by taking a closer look at what happens to students' competence beliefs when students' utility value beliefs are manipulated (cf., Pintrich, 2003). Thereby, the question of the direction of the association between students' competence beliefs and value beliefs is addressed (cf., Wigfield et al., 2009).

### 3.2 Sample and design of the empirical studies

The three empirical studies conducted in the context of this dissertation were based on data from the **intervention project "Motivation in Mathematics" (MoMa)** conducted at the University of Tübingen during the school year of 2012/2013. The data were collected from 82 mathematics classes in 25 academic track schools in Baden-Württemberg, Germany. Academic track schools are the most advanced of the three types of secondary school in Germany and prepare students for university study. Classes were randomly assigned to either one of two intervention conditions or a control condition. This cluster-randomized study design was similar to students' genuine classroom experience and allowed the development of an intervention which involved activating students, for example, through discussions with partners, while diminishing the risk of diffusion effects (Craven et al., 2001). The interventions in this study were based on results of prior research indicating that relevance interventions work best when students are confident they are able to complete the target task successfully or succeed in the subject (e.g., Durik et al., 2015) and that they are particularly effective when combining the provision and the self-generation of relevance information (Canning & Harackiewicz, 2015). The interventions in this study consisted of a psychoeducational presentation and an individual writing activity differing for each condition. The effectiveness of the two writing activities—one adapted from prior intervention research ("text", e.g., Hulleman & Harackiewicz, 2009), one newly developed ("quotations")—was systematically compared.

Students' competence beliefs and value beliefs were assessed before the interventions and again six weeks and five months after the interventions using students' self-reports and the comprehensive value instrument developed by Gaspard, Dicke, Flunger, Schreier et al. (2015). Students' perceptions of the classroom context (teaching strategy "emphasis on the practical application of the mathematics topic being taught" and the value classmates attribute to learning mathematics) were assessed at each assessment point. Teachers reported on their teaching strategies "introducing new mathematics topics with examples from everyday life" and "demonstrating links between mathematics and other academic subjects" at the pretest, and they rated students' effort at each of the three assessment points. Students' achievement in mathematics was assessed according to their scores on a state-wide standardized mathematics test before the intervention and their scores on a short standardized mathematics test measuring students' fluency of solving typical math operations (Schmidt, Ennemoser, & Krajewski, 2013) five months after the intervention. Furthermore, students' essays produced during the intervention were collected and coded to measure their responsiveness.

### 3.3 Research questions of the empirical studies

In **Study 1**, entitled *Der Wert der Mathematik im Klassenzimmer—Die Bedeutung relevanzbezogener Unterrichtsmerkmale für die Wertüberzeugungen der Schülerinnen und Schüler [The value of mathematics in the classroom: The importance of a relevance-oriented learning environment for students' value beliefs]*, the effect of instructional strategies and the value classmates attribute to learning mathematics on students' mathematics-related value beliefs was investigated. Results of this study thereby contribute to filling a gap in research framed in EVT on the role of students' socializers (teachers, peers) and their beliefs and behaviors for students' development of all four value beliefs. Unlike prior studies of students' motivation in the classroom, in this dissertation the social complexity of classroom was taken into account by simultaneously investigating teacher- and peer-related influences, and changes in value beliefs were investigated using a longitudinal design (see Frenzel et al., 2010, for an exception). The following research questions were investigated in Study 1:

1) How are relevance-oriented teaching strategies in mathematics (stressing practical applicability, introducing new topics with everyday examples, demonstrating links with other academic subjects) and students' perception of the value their classmates attribute to learning mathematics associated with students' mathematics-related value beliefs (intrinsic, attainment, utility values, and cost)?
2) Do relevance-oriented teaching strategies and students' perception of the value their classmates attribute to learning mathematics lead to a change in students' mathematics-related value beliefs after six months?

To address the first research question pretest data from the MoMa intervention study were analyzed in multiple linear regression models distinguishing between the individual level and the class level. To address the second research question, data from the first and the last measurement points (six months after the pretest) of the MoMa intervention were examined in multiple hierarchical linear regression models controlling for students' initial value beliefs as well as for the intervention.

In **Study 2**, entitled *Short intervention, sustained effects: Promoting students' mathematics-related competence beliefs, effort, and achievement*, investigation was made into the effects of the MoMa relevance interventions on students' competence beliefs, effort, and achievement, thereby extending relevance intervention research to secondary school classrooms in Germany. Unlike in previous relevance intervention studies, the effectiveness of two intervention approaches (one previously established and one newly developed) was systematically compared. Through implementation at the class level, the interventions were first adapted to students' genuine classroom setting. The short- and long-term effects of the relevance interventions on a broad range of previously neglected outcomes including students' motivation, behavior, and achievement in mathematics were analyzed. In a previous investigation, the interventions had been found to foster students' value beliefs (Gaspard, Dicke, Flunger, Brisson et al., 2015). Study 2 addresses the following research questions:

1) How do two relevance interventions (writing a text or evaluating quotations about the relevance of mathematics) influence students' mathematics-related competence beliefs (self-concept, homework self-efficacy) and their effort to learn mathematics as rated by teachers six weeks after the intervention?
2) Are the intervention effects stable?
3) Do the interventions promote students' test-based achievement in mathematics five months after the intervention?

To answer the three research questions, separate multiple linear hierarchical regression models were run for each of the outcomes six weeks and five months after the interventions. Two dummies representing the intervention conditions at the class level were simultaneously regressed on students' competence beliefs, effort, and achievement, controlling for students' initial values.

In **Study 3**, entitled *Who sticks to the instructions—and does it matter? Antecedents and effects of students' fidelity to a classroom-based relevance intervention*, investigation was made into the processes underlying the effects of the MoMa relevance interventions. Little is known about the mechanisms through which relevance interventions work or do not work. Knowledge about the characteristics of students who respond well and those who respond less well to relevance interventions is needed to find ways to optimize relevance interventions so as to reach a maximum number of students. By systematically analyzing the criteria "positive argumentation", "personal connections", and "in-depth reflection", the third study also makes a unique contribution to understanding the elements through which a change in students' perception of relevance can be triggered. In Study 3 the following research questions were addressed:

1) How did students respond to the writing tasks in the MoMa relevance interventions?
2) Which individual student characteristics and classroom-related perceptions predicted students' responsiveness to the writing tasks?
3) How does the degree of students' responsiveness relate to the effects of the interventions on students' utility value beliefs six weeks and five months after the intervention?

To address these research questions, students' essays produced during the MoMa interventions were coded according to the three fidelity criteria which were first investigated descriptively and then combined into a responsiveness index. Students' individual characteristics and classroom-related perceptions at the pretest were regressed on the responsiveness index using multiple linear regression models. Finally, complier-average causal effects analyses (e.g., Sagarin et al., 2014) were conducted to compare the effects of the relevance interventions on the utility value beliefs of the responsive students and the nonresponsive students six weeks and five months after the intervention.

# References


Acatech & Körber-Stifung (2014). *MINT-Nachwuchsbarometer 2014*. München/Hamburg.

Acatech & Körber-Stifung (2015). *MINT-Nachwuchsbarometer 2015*. München/Hamburg.

Allen, V. L., & Feldman, R. S. (1976). Studies on the role of tutor. In V. L. Allen (Ed.), *Children as teachers. Theory and research on tutoring* (pp. 113–130). New York: Academic Press.

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. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1), 1–22. https://doi.org/10.1111/j.1741-5446.2006.00241.x

Brophy, J. E. (1999). Teaching. Lausanne, Switzerland: PCL.

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility value information. *Motivation Science, 1*(1), 47–71. https://doi.org/10.1037/mot0000015

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Craven, R. G., Marsh, H. W., Debus, R. L., & Jayasinghe, U. (2001). Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. *Journal of Educational Psychology, 93*(3), 639–645. https://doi.org/10.1037/0022-0663.93.3.639

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

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). New York: Freeman.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Maciver, D. (1993). Development during adolescence—The impact of stage-environment fit on young adolescents experiences in schools and in families. *American Psychologist, 48*(2), 90–101. https://doi.org/10.1037//0003-066x.48.2.90

Eccles, J. S., Vida, M. N., & Barber, B. (2004). The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *Journal of Early Adolescence, 24*(1), 63–77. https://doi.org/10.1177/0272431603260919

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment: Monitoring behavior in daily life settings. *European Journal of Psychological Assessment, 23*(4), 206–213. https://doi.org/10.1027/1015-5759.23.4.206

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*(9), 1226–1240. https://doi.org/10.1037 /dev0000028

Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B. M., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open, 2*(2), 1–14. https://doi.org /10.1177/2332858416649168

Häfner, I., Flunger, B., Dicke, A.-L., Gaspard, H., Brisson, B. M., Nagengast, B., & Trautwein, U. (2017). Robin Hood effects on motivation in math: Family interest moderates the effects of relevance interventions. *Developmental Psychology, 53*(8), 1522–1539. https://doi.org /10.1037 /dev0000337

Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytical review. *Journal of Clinical Child Psychology, 27*(4), 423–433.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*(5), 745–765. https://doi.org/10.1037 /pspp0000075

Harackiewicz, J. M., Hulleman, C. S., Rozek, C., Katz-Wise, S. L., & Hyde, J. S. (2010, March). *Parents' understanding of the utility value of STEM courses for high school students.* Paper presented at the 13th Biennial SRA Conference (Society for Research on Adolescence). Philadelphia, PA.

Hoogerheide, V., van Wermeskerken, M., Loyens, S. M. M., & van Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction, 44*, 22–30. https://doi.org/10.1016/j.learninstruc.2016.02.004

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*(1), 88–110. https://doi.org/10.1080/19345740802539325

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*(4), 880–895. https://doi.org/10.1037/A0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*(3), 387–404. https://doi.org/10.1037/edu0000146

Husman, J., Nelson, K., & Cheng, K. (2017, August). *Intervening to promote engineering student success in technical, nonmajor, required barrier courses.* Poster presented at the 17th Biennial EARLI Conference (European Association for Research on Learning and Instruction). Tampere, Finland.

Institut der deutschen Wirtschaft. (2017). MINT-Frühjahrsreport 2017: *MINT-Bildung: Wachstum für die Wirtschaft, Chancen für den Einzelnen [Education in STEM: Economic growth, opportunities for the individual].* http://www.mintzukunftschaffen.de /uploads/media/FINAL_MINT-Fruehjahrsreport_2017.pdf

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*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588–600. https://doi.org/10.1037/a0019682

Jo, B. (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics, 27*(4), 385–409. https://doi.org/10.3102/10769986027004385

Juvonen, J., Espinoza, G., & Knifsend, C. (2012). The role of peer relationships in student academic and extracurricular engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 387–492). New York: Springer.

Karabenick, S. A., Albrecht, J., & Rausch, N. (2017, August). *Examining the self-generation of subjective task value: A mixed-methods investigation.* Paper presented at the 17th Biennial EARLI Conference (European Association for Research on Learning and Instruction). Tampere, Finland.

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Development, 78*(4), 1186–1203. https://doi.org/10.1111/j.1467-8624.2007.01060.x

Kindermann, T. A., McCollam, T. L., & Gibson, E. (1996). Peer networks and students' classroom engagement during childhood and adolescence. In J. Juvonen & K. R. Wentzel (Eds.), *Social motivation: Understanding children's school adjustment* (pp. 279–312). Cambridge, England: Cambridge University Press.

Klein, F., & Schimmack, R. (1907). *Vorträge über den Mathematischen Unterricht an den Höheren Schulen. Mathematische Vorlesungen an der Universität Göttingen*. Leipzig: B. G. Teubner.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231–251. https://doi.org/10.1007/s10984-006-9015-7

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning: Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133–149. https://doi.org/10.1016/j.edurev.2013.02.002

Lazarides, R., & Ittel, A. (2012). Instructional quality and attitudes toward mathematics: Do self-concept and interest differ across students' patterns of perceived instructional quality in mathematics classrooms? *Child Development Research, 2012*, 1–11. https://doi.org/10.1155/2012/813920

Lazarides, R., & Rubach, C. (2017). Instructional characteristics in mathematics classrooms: Relationships to achievement goal orientation and student engagement. *Mathematics Education Research Journal, 29*(2), 201–217. https://doi.org/10.1007/s13394-017-0196-4

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423–458. https://doi.org/10.3102/00346543066004423

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312–320. https://doi.org/10.1037/0022-0663.90.2.312

Murrah, W. M., Kosovich, J. J., & Hulleman, C. S. (2017). A framework for incorporating intervention fidelity in educational evaluation studies. In G. Roberts, S. Vaughn, T. Beretvas, & V. Wong (Eds.), *Treatment fidelity in studies of educational intervention* (pp. 39–60). New York: Routledge.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*(8), 1058–1066. https://doi.org/10.1177/0956797611415540

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181–206. https://doi.org/10.1207/s15326985ep4103_4

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). New York: Springer.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Ramani, G. B., Zippert, E., & Daubert, E. (2016). The influence of same- and cross-age peers on children's literacy and mathematical development. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of social influences in school contexts. Social-emotional, motivation, and cognitive outcomes* (pp. 96–112). New York, London: Routledge Taylor & Francis Group.

Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H.,. . . Wang, C. K. J. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion, 38*(1), 93–110. https://doi.org/10.1007/s11031-013-9367-0

Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E., & Köller, O. (Eds.). (2016). *PISA 2015: Eine Studie zwischen Kontinuität und Innovation. Münster: Waxmann*. Retrieved from http://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtband_und_Zusammenfassung_2012/PISA_2015_eBook.pdf

Renn, O., Duddeck, H., Menzel, R., Holtfrerich, C.-L., Luca, K., Fischer, W.,. . . Pfenning, U. (2012). *Stellungnahmen und Empfehlungen zur MINT-Bildung in Deutschland auf der Basis einer europäischen Vergleichsstudie*. Berlin.

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*(4), 327–362. https://doi.org/10.1007/s10648-005-8137-2

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM Motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist, 51*(2), 146–163. https://doi.org/10.1080/00461520.2016.1154792

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78.

Sarbin, T. R. (1976). Cross-age tutoring and social identity. In V. L. Allen (Ed.), *Children as teachers. Theory and research on tutoring* (pp. 27–40). New York: Academic Press.

Schmiedek, F. (2016). Methods and Designs. In T. Strobach & J. Karbach (Eds.), *Cognitive training. An overview of features and applications* (pp. 9–18). Cham: Springer International.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*(2), 149–174. https://doi.org/10.2307/1170234

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly, 23*(1), 7–25.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Belmont, CA: Wadsworth Cengage Learning.

Topping, K. (2000). *Tutoring (5th ed.). Educational Practice Series.* Lausanne, Switzerland: PCL.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. https://doi.org/10.1037/A0027470

Turner, J. C., Warzon, K. B., & Christensen, A. (2011). Motivating mathematics learning: Changes in teachers' practices and beliefs during a nine-month collaboration. *American Educational Research Journal, 48*(3), 718–762. https://doi.org/10.3102/0002831210385103

Turner, R. H. (2006). Role theory. In J. H. Turner (Ed.), *Handbooks of sociology and social research: Handbook of sociological theory* (pp. 233–254). Dordrecht: Springer.

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology, 48*(6), 1643–1657. https://doi.org/10.1037/A0027247

Wentzel, K. R. (2005). Peer relationships, motivation, and academic performance at school. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 279–296). New York and London: Guilford.

Wentzel, K. R., & Brophy, J. E. (2014). *Motivating students to learn (4th ed.).* New York: Routledge.

Wigfield, A., Tonks, S., & Klauda, S. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–75). New York: Routledge.

Woolley, M. E., Rose, R. A., Orthner, D. K., Akos, P. T., & Jones-Sanpei, H. (2013). Advancing academic achievement through career relevance in the middle grades: A longitudinal evaluation of CareerStart. *American Educational Research Journal, 50*(6), 1309–1335. https://doi.org/10.3102/0002831213488818

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301. https://doi.org/10.3102/0034654311405999

# EMPIRICAL STUDIES

# 4
# Study 1

**Der Wert der Mathematik im Klassenzimmer:**

**Die Bedeutung relevanzbezogener Unterrichtsmerkmale für die Wertüberzeugungen der Schülerinnen und Schüler**

[The value of mathematics in the classroom:

The importance of a relevance-oriented learning environment for students' value beliefs]

Schreier, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Lüdtke, O., Nagengast, B., & Trautwein, U. (2014). Der Wert der Mathematik im Klassenzimmer. Die Bedeutung relevanzbezogener Unterrichtsmerkmale für die Wertüberzeugungen der Schülerinnen und Schüler. *Zeitschrift für Erziehungswissenschaft, 17* (2), 225–255. DOI: 10.1007/s11618-014-0537-y

## Zusammenfassung

Mangelnde Schülermotivation stellt im Mathematikunterricht der Sekundarstufe eine große Herausforderung dar. Lehrkräfte, Mitschülerinnen und Mitschüler beeinflussen laut Erwartungs-Wert-Modell der Leistungsmotivation (Eccles et al., 1983) die Wertüberzeugungen. Mit Daten von 1868 Neuntklässlerinnen und Neuntklässlern und ihren 72 Mathematiklehrkräften wurde der Zusammenhang zwischen relevanzbezogenen Unterrichtsmerkmalen und den Wertüberzeugungen der Schülerinnen und Schüler in Mathematik untersucht. Lineare Regressionsanalysen auf zwei Ebenen zeigten, dass die aus Schülersicht berichtete Praxisorientierung im Mathematikunterricht vor allem auf individueller Ebene und die aus Schülersicht wahrgenommene Wertschätzung des Fachs im Klassenverband vor allem auf Klassenebene mit intrinsischem Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugungen (Schülerbericht) assoziiert waren. Im Verlauf von sechs Monaten verstärkten sich diese Effekte nur auf Schülerebene. Die aus Lehrersicht erfasste Demonstration von Sachverbindungen sagte Wichtigkeitsüberzeugungen positiv vorher und die Themeneinführung mit Alltagsbeispielen führte zur Abnahme der Kostenüberzeugungen der Schülerinnen und Schüler innerhalb von sechs Monaten.


**Schlüsselwörter:** Wertüberzeugungen · Motivation · Mathematikunterricht · Lehrkraft · Mitschüler/innen

## Abstract

The lack of secondary school students' motivation in mathematics lessons poses a great challenge. According to the expectancy value model of achievement motivation (Eccles et al., 1983), teachers and classmates influence students' value beliefs. Using data of 1868 ninth grade students and their 72 mathematics teachers, this study was designed to assess the association of several indicators of relevance-oriented teaching strategies and of students' perception of the value their classmates attributed to learning mathematics with students' own mathematics-related value beliefs. Two-level linear regression analyses showed that the student-reported strategy "stressing the practical applicability of mathematics" predicted students' intrinsic, attainment, utility, and cost values mainly at the individual level; students' perception of the value their classmates attributed to learning mathematics was associated with students' own value perceptions more strongly at the class level. Over the course of six months, these effects increased at the individual level but not at the class level. Teacher-reported "demonstrating links between mathematics and other academic subjects" was positively associated with students' attainment and cost values. Teacher-reported use of daily life examples led to a decrease in students' cost value over the course of six months.


**Keywords:** value beliefs · motivation · teaching strategies · classroom environment · classmates

## Einleitung

„Wozu müssen wir dieses Zeug lernen?“—Mathematiklehrkräfte der Sekundarstufe sind vermutlich vertraut mit dieser Frage. Gerade im Jugendalter ist es Mathematiklehrkräften ein großes Anliegen, durch ihren Unterricht für eine hohe Arbeitsmotivation der Schülerinnen und Schüler zu sorgen (vgl. Pierce und Stacey 2006). Die Lernumgebung ist in der Tat zentral für die motivationale Entwicklung von Schülerinnen und Schülern (Eccles und Wigfield 2002; Pekrun 2006). Im Hinblick auf die Wertüberzeugungen in naturwissenschaftlichen Fächern deuten empirische Forschungsarbeiten darauf hin, dass es für Wichtigkeits- und Nützlichkeitsüberzeugung förderlich ist, wenn im Unterricht die persönliche Relevanz der Lerninhalte für die Schülerinnen und Schüler thematisiert wird (z. B. Hulleman und Harackiewicz 2009; Wang 2012). Für die Entwicklung von Schülerinteresse ist außerdem die aus Schülersicht wahrgenommene Wertschätzung im Klassenverband gegenüber dem Fach Mathematik bedeutsam (z. B. Frenzel et al. 2010). Unklar ist jedoch, welche Unterrichtsmerkmale die stärkste Rolle für die vier Wertkomponenten intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung von Schülerinnen und Schülern (vgl. Erwartungs-Wert-Modell nach Eccles et al. 1983) spielen. So können Unterrichtsstrategien zur Vermittlung der Relevanz mathematischer Lerninhalte—wie beispielsweise die Erarbeitung neuer Themen durch Alltagsbeispiele (vgl. Vorhölter 2009) oder die Verbindung der Mathematik mit den Inhalten anderer Fächer (vgl. Michelsen und Sriraman 2009) – oder aber die wahrgenommene Wertschätzung im Klassenverband unterschiedlich stark mit den verschiedenen Komponenten der Wertüberzeugungen von Schülerinnen und Schüler in Mathematik zusammenhängen. Unter Berücksichtigung der Schüler- und der Lehrerperspektive wurde in der vorliegenden Studie erstmals differenziell der Zusammenhang zwischen relevanzbezogenen Unterrichtsmerkmalen und den vier Wertüberzeugungen von Sekundarschülerinnen und -schülern in Mathematik untersucht.

## Theoretischer Hintergrund

### Zur motivationalen Bedeutung der Wertüberzeugungen in Mathematik

Schülermotivation wird häufig im Rahmen der Erwartungs-Wert-Theorie leistungsbezogener Verhaltensweisen und Wahlentscheidungen nach Eccles et al. (1983) erfasst. Eccles‘ Erwartungs-Wert-Modell geht unter anderem davon aus, dass vier subjektive Wertkomponenten—intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung—das leistungsbezogene Verhalten von Schülerinnen und Schülern beeinflussen (Eccles und Wigfield 2002). Bezogen auf das Fach Mathematik werden die Wertüberzeugungen wie folgt definiert: Intrinsischer Wert bezeichnet die Freude, die Schülerinnen und Schüler bei der Beschäftigung mit mathematischen Aufgaben empfinden. Wichtigkeit bezieht sich darauf, wie viel es Schülerinnen und Schülern bedeutet, gut in Mathematik zu sein, und für wie zentral sie Mathematikkenntnisse für ihre persönliche Identität halten. Diese ersten beiden Wertkomponenten weisen beträchtliche Überschneidungen mit dem Konstrukt des individuellen bzw. persönlichen Interesses auf (vgl. z. B. Schiefele 2009; Wigfield und Cambria 2010). Weiter schreiben Schülerinnen und Schüler der Mathematik Nützlichkeit zu, wenn sie glauben, dass mathematisches Wissen ihnen dabei helfen kann, ihre individuellen Ziele zu erreichen. Die Wertkomponente Kosten bezieht sich auf sämtliche negative Aspekte, die für Schülerinnen und Schüler mit der Beschäftigung mit Mathematikaufgaben einhergehen, wie beispielsweise der Zeitverlust für andere Aktivitäten oder die erwartete Anstrengung, die aufgebracht werden muss, um eine Aufgabe erfolgreich zu lösen.

Die Bedeutung dieser vier Wertkomponenten für leistungsbezogenes Schülerverhalten konnte nicht nur für Mathematik sondern für eine große Anzahl von Schulfächern empirisch nachgewiesen werden. So bestehen für intrinsischen Wert, Wichtigkeits- und Nützlichkeitsüberzeugung von Schülerinnen und Schülern positive und für die Kostenüberzeugung negative Zusammenhänge mit Anstrengungsbereitschaft, Durchhaltevermögen, Konzentration, Leistung und kognitiver Auseinandersetzung mit Lerninhalten sowie mit Kurswahlentscheidungen (im Überblick z. B. Roeser et al. 2000; Vansteenkiste et al. 2004; Wigfield et al. 2009). Gleichzeitig weisen Forschungsarbeiten darauf hin, dass Wertüberzeugungen und Interesse von Schülerinnen und Schülern im Sekundarschulalter abnehmen (im Überblick: Wigfield und Eccles 2002; Wigfield et al. 2006). In naturwissenschaftlichen Fächern allgemein und in speziell in Mathematik gilt diese Abnahme als besonders drastisch (z. B. Jacobs et al. 2002; Watt 2004; Chouinard und Roy 2008; Krapp und Prenzel 2011). Vor dem Hintergrund dieser Entwicklungen interessiert uns die Frage: Was charakterisiert eine Lernumgebung, in der Schülerinnen und Schüler der Mathematik einen hohen Wert beimessen? Welche Rolle spielen die Lehrkräfte, ihre Unterrichtsstrategien und die Klassenkameradinnen und -kameraden?

**Wertüberzeugungen und die Rolle von Bezugspersonen im Klassenzimmer**

Wie Überblickswerke zeigen (z. B. Wigfield et al. 2006), untersuchte bereits eine Reihe von Studien die Auswirkungen des Lehrer- und Mitschülerverhaltens auf die Schülermotivation, jedoch auf Grundlage unterschiedlicher Motivationstheorien. Etliche dieser Sammelwerke betonen, wie wichtig es zur Unterstützung von Schülermotivation ist, dass die Lehrkraft im Unterricht Bezüge zwischen dem Lernmaterial und der Lebenswelt der Schülerinnen und Schüler herstellt und dazu Aufgaben auswählt, die authentisch und inhaltlich bedeutsam für die Lernenden sind (z. B. Fredricks et al. 2004; Pianta et al. 2012; Wigfield et al. 2006). Weiter wird in Überblicksbeiträgen resümiert, dass die Leistungsmotivation von Schülerinnen und Schülern von der in der Peergroup wahrgenommenen schulischen Motivation abhängt (z. B. Fredricks et al. 2004; Juvonen et al. 2012). Die Erwartungs-Wert-Theorie nimmt an, dass auch die subjektiven Wertüberzeugungen von Schülerinnen und Schülern vom Verhalten ihrer Bezugspersonen beeinflusst werden (Eccles und Wigfield 2002).

**Lehrkraft und ihre Unterrichtsstrategien.** Dass die Lehrkraft im Klassenzimmer eine wichtige Bezugsperson darstellt, dürfte unumstritten sein. Durch ihre Unterrichtsführung initiiert und unterstützt sie nicht nur Lernprozesse, sondern vermittelt auch die Wertentwicklung von Schülerinnen und Schülern: Laut Pekrun (2006) kann die Lehrkraft positive Lernemotionen fördern, indem sie den Lernenden klar macht, dass sich Anstrengung beim

Wissenserwerb lohnt, weil das Gelernte „wertvoll“ ist. Eine solche „Wertinduktion“ (Pekrun 2006, S. 334) kann nicht nur durch direkte verbale Mitteilungen erfolgen, sondern auch durch bestimmte Herangehensweisen im Unterricht. Diese Ansicht wird auch in der mathematik-didaktischen Forschung vertreten (z. B. Freudenthal 1968; Blum und Niss 1991).

Einige Studien (z. B. Frenzel et al. 2010) beschäftigten sich bereits mit der Rolle der Lehrkraft und ihrer Unterrichtsstrategien für die Entwicklung von Interesse und von einzelnen Wertkomponenten wie intrinsischem Wert, Wichtigkeitsüberzeugung und Leistungsangst—einem der Kostenüberzeugung ähnlichen Konstrukt—im Mathematikunterricht der Sekundarstufe. Dabei konnten beispielsweise Enthusiasmus und Erwartungshaltung der Lehrkraft, Förderung des kooperativen Lernens und der Schülerautonomie, gut strukturierter und kognitiv aktivierender Unterricht sowie individuelle Unterstützung als Prädiktoren identifiziert werden (z. B. Daniels 2008; Frenzel et al. 2010; Kunter et al. 2007; Kunter und Voss 2013; Wang 2012). Empirische Studien legen außerdem nahe, wie bedeutsam es für Schülermotivation ist, wenn die Lehrkraft den Schülerinnen und Schülern dabei hilft, den Lernprozess als relevant für eigene Ziele und Interessen zu erkennen. So steigen positive Lernemotionen sowie behaviorale und kognitive Auseinandersetzung bei Sekundarschülerinnen und -schülern, je mehr Relevanzförderung sie im Unterricht wahrnehmen (Assor et al. 2002). Weiter ist die wahrgenommene Relevanz des Lehrplans positiv mit Änderungen in der allgemeinen Wertschätzung der Schule während der Sekundarstufe assoziiert (Roeser et al. 1998).

Obwohl der Relevanzorientierung im Unterricht gerade in den Naturwissenschaften (z. B. Osborne und Dillon 2008) bzw. im Fach Mathematik eine besondere Rolle zur Motivations- und Interessensförderung zugeschrieben wird (z. B. Blum und Niss 1991; Krapp 1998; Turner und Meyer 2009), untersuchten bisher nur wenige Studien entsprechende Zusammenhänge mit den Wertüberzeugungen von Schülerinnen und Schülern in Mathematik. Studien von Wang (2012) und Willems (2011) zeigten, dass die wahrgenommene Bedeutsamkeit des Lehrplans und der Unterrichtsmethoden in Mathematik Interesse, intrinsischen Wert und Wichtigkeitsüberzeugung von Sekundarschülerinnen und -schülern positiv beeinflusst. Die individuell wahrgenommene Bedeutsamkeit spielt dabei eine deutlich stärkere Rolle für Schülerinteresse als die in der Klasse geteilte Wahrnehmung (Willems 2011). Hulleman und Harackiewicz (2009) fanden in einer Interventionsstudie zudem, dass Sekundarschülerinnen und -schüler Naturwissenschaften für interessanter und nützlicher hielten, wenn sie in Schreibaufgaben die persönliche Bedeutung des Gelernten erarbeiteten.

Bislang fehlen allerdings Studien, die den Zusammenhang zwischen Relevanzorientierung und Wertüberzeugungen in Mathematik mit ausreichender Differenzierung in Bezug auf die vier Wertkomponenten und die Unterrichtsmerkmale untersuchten. So wurden bisher selten alle sondern meist nur einzelne Wertkomponenten analysiert—und diese zudem auf sehr unterschiedliche Art erfasst. Dies spiegelt die potentielle Vielfältigkeit von intrinsischem Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugungen wider, welcher nur durch den Einsatz von umfassenderen Erhebungsinstrumenten Rechnung getragen werden kann (vgl. Trautwein et al. 2013).

Was die Unterrichtsmerkmale betrifft, wurde bisher nur am Rande geklärt, mit welchen konkreten Strategien Lehrkräfte den Wert der Mathematik erfolgreich aufzeigen können. Freudenthal (1968) postulierte, dass Mathematikunterricht dann „nützlich" ist, wenn Schülerinnen und Schüler lernen, wie sie ihre Kenntnisse im Alltag einsetzen können. Hierbei sollte ein mathematisches Problem ausgehend von einer Alltagssituation definiert werden. In qualitativen Studien wurde gezeigt, dass solche sogenannten Modellierungsaufgaben Schülerinnen und Schülern dabei helfen, den Sinn der Mathematik zu erkennen (Vorhölter 2009). Mathematiklehrkräfte geben außerdem an, Aufgaben mit Alltagsbezug gezielt im Unterricht einzusetzen, um die Einstellung der Lernenden zur Mathematik zu verbessern (Pierce und Stacey 2006). Des Weiteren konnten Rakoczy et al. (2008) die Bedeutung des wahrgenommenen Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation empirisch nachweisen. Die interdisziplinäre Verknüpfung der Mathematik mit anderen Fachbereichen stellt eine weitere Unterrichtsstrategie dar, die die Anwendbarkeit und dadurch die Erkenntnis der Relevanz der Mathematik fördern könnte (z. B. Blum und Niss 1991). Eine Umfrage von Michelsen und Sriraman (2009) liefert erste Hinweise darauf, dass Sekundarschülerinnen und -schüler Mathematik interessanter finden, wenn im Unterricht mathematische Inhalte mit den Inhalten anderer Fachbereiche in Verbindung gebracht werden.

Diese Arbeiten deuten an, dass die Unterrichtsstrategien „Themeneinführung mit Alltagsbeispielen" und „Demonstration von Sachverbindungen" hilfreich zur Wertvermittlung in Mathematik sein könnten. Um dies differenziert zu überprüfen, könnte der Einbezug der Lehrerperspektive einen großen Mehrwert darstellen—selbst wenn sich die Lehrersicht auf den Unterricht häufig als weniger prädiktiv für Schülerverhalten erweist als die Sicht der Schülerinnen und Schüler selbst (vgl. Clausen 2002; Kunter und Baumert 2006). Zum Beispiel sind Lehrkräfte in Bezug auf die Aufgabenart und die dazugehörigen Erklärungen in der Lage, eine differenziertere Perspektive einzunehmen als ihre Schülerinnen und Schüler (Kunter und Baumert 2006).

**Klassenkameradinnen und -kameraden und ihre Wertschätzung gegenüber Mathematik.** Im Unterrichtsgeschehen gelten neben der Lehrkraft auch die Klassenkameradinnen und -kameraden als wichtige Bezugspersonen von Schülerinnen und Schülern. Die Sekundarstufe stellt dabei einen Abschnitt in der Schulkarriere dar, in der sich die Beziehungsinteressen der Heranwachsenden immer stärker auf Gleichaltrige verlagern; gleichzeitig bilden die Jugendlichen zunehmend ihre eigenen Wertvorstellungen aus und achten bei Gleichaltrigen weniger auf deren Verhalten als auf deren Charaktereigenschaften und Wertvorstellungen (Parker et al. 2006). Im Klassenzimmer erfahren Schülerinnen und Schüler die Einstellungen ihrer Mitschülerinnen und -schüler im täglichen Umgang miteinander (Ryan 2000). Gerade in einem Schulsystem wie dem deutschen, in welchem die Schülerinnen und Schüler jahrelang im gleichen Klassenverband zusammen lernen, können Klassenkameradinnen und -kameraden ihre gegenseitigen schulischen Werte besonders intensiv erleben (vgl. Frenzel et al. 2010). Durch die zunehmende Orientierung auf Peers und deren Werte im Jugendalter ist zu vermuten, dass nicht nur Aspekte der inhaltlichen Ausrichtung des Mathematikunterrichts, sondern auch sozial-normative Aspekte wie die wahrgenommene Wertschätzung des Fachs Mathematik im Klassenverband die Wertüberzeugungen von Schülerinnen und Schülern beeinflussen können (vgl. auch Eccles et al. 1993).

Es gibt bereits einige empirische Belege für die Bedeutung der im sozialen Umfeld wahrgenommenen motivationalen Haltung für die individuelle Motivation von Schülerinnen und Schülern der Sekundarstufe. So beeinflussen Freunde und Freundinnen zu Beginn der Sekundarstufe im Laufe eines Schuljahres ihre gegenseitige schulische Motivation zwar nur leicht, in der Tendenz aber verstärkend (Altermatt und Pomerantz 2003; Kindermann 2007). Ryan (2001) stellte zudem fest, dass im Freundeskreis wahrgenommene Wertüberzeugungen bezüglich Schule im Zusammenhang mit der Entwicklung von schulischem Interesse steht, nicht aber mit Veränderungen in der Wichtigkeit und Nützlichkeit, die der Schule zugeschrieben wird. Im Fach Naturwissenschaften stellten sich sowohl die Werte der besten Freundinnen und Freunde als auch die wahrgenommenen leistungsbezogenen Normen im Klassenverband als positive Prädiktoren für die individuellen Leistungsziele und sozialen Werte von Sekundarschülerinnen und -schülern heraus (Nelson und DeBacker 2008). Studien von Frenzel et al. (2007; 2010) ergänzen diesen Befund um die Bedeutung des Werteklimas im Klassenverband als Kontextmerkmal für die klassenspezifische motivationale Entwicklung von Schülerinnen und Schülern. So konnten die Autoren zeigen, dass eine höhere wahrgenommene Wertschätzung der Mathematik im Klassenverband sowohl auf individueller Ebene als auch auf Klassenebene mit mehr Interesse und positiveren Emotionen im Mathematikunterricht einhergeht.

Insgesamt deuten diese Studien an, dass sich die subjektiv wahrgenommenen Werte im Freundeskreis und im Klassenverband potentiell nicht nur auf Schülerebene sondern auch aggregiert als Indikator des Klassenwerteklimas auf die Schülermotivation auswirken können. Die Bedeutung der subjektiv wahrgenommenen Wertschätzung der Mathematik im Klassenverband für die individuellen und die klassenspezifischen Wertüberzeugungen von Schülerinnen und Schülern in Mathematik wurde bislang jedoch noch nicht untersucht.

## Ziele der vorliegenden Studie

Bisherige Forschungsarbeiten legen die Vermutung nahe, dass Mathematikunterricht, in welchem die Lehrkraft die Relevanz der Lerninhalte thematisiert, mit den Wertüberzeugungen der Schülerinnen und Schüler zusammenhängen könnte (vgl. Hulleman und Harackiewicz 2009; Wang 2012). Auch die wahrgenommene Wertschätzung der Mathematik im Klassenverband könnte mit den Wertüberzeugungen in Verbindung stehen (vgl. Nelson und DeBacker 2008; Frenzel et al. 2010). Eine umfassende Untersuchung dieser Zusammenhänge fehlt bisher. Es ist unklar, welche konkreten Unterrichtsstrategien sich zur Vermittlung der Relevanz mathematischer Lerninhalte eignen und welche Rolle die wahrgenommene Wertschätzung im Klassenverband für die Wertüberzeugungen der Schülerinnen und Schüler in Mathematik spielt. Die vorliegende Studie zielt darauf ab, diese Forschungslücken zu füllen. Im Bereich der Unterrichtsstrategien wurden die Themeneinführung mit Alltagsbeispielen (vgl. Vorhölter 2009) und die Demonstration von Verbindungen zwischen Mathematik und den Inhalten anderer Fächer (vgl. Michelsen und Sriraman 2009) als Prädiktoren untersucht. Da Lehrkräfte bei der Aufgabenart und bei Erklärungen im Unterricht stärker differenzieren als Schülerinnen und Schüler (vgl. Kunter und Baumert 2006), wurden diese Prädiktoren aus Lehrersicht erhoben. Die aus Schülersicht beurteilte Praxisorientierung im Mathematikunterricht stellte einen weiteren Prädiktor dar. Dieser gibt an, wie stark die Schülerinnen und Schüler subjektiv wahrnahmen, dass die Lehrkraft im Unterricht die Anwendbarkeit der Mathematik thematisierte (vgl. Wang 2012). Zuletzt wurde die aus Schülersicht wahrgenommene Wertschätzung der Mathematik im Klassenverband als Prädiktor erfasst. Für eine möglichst aussagekräftige Erfassung der abhängigen Variablen intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugungen wurde ein differenziertes Erhebungsinstrument eingesetzt (vgl. Trautwein et al. 2013).

So lauten die Forschungsfragen: Wie hängen Merkmale des relevanzorientierten Mathematikunterrichts (erfasst durch die Praxisorientierung, die Themeneinführung mit Alltagsbeispielen und die Demonstration von Sachverbindungen) sowie die wahrgenommene Wertschätzung der Mathematik im Klassenverband mit intrinsischem Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugungen von Schülerinnen und Schülern in Mathematik zusammen? Führen diese Unterrichtsmerkmale zu Veränderungen in den Wertüberzeugungen innerhalb von sechs Monaten?

Da die untersuchten Faktoren im realen Unterrichtsgeschehen gebündelt auftreten, sollen die Zusammenhänge dieser Unterrichtsmerkmale mit den Wertüberzeugungen und deren Entwicklung zunächst relativ zueinander analysiert und die univariaten Assoziationen im Anschluss überprüft werden. In Anlehnung an bisherige Forschungsarbeiten nehmen wir positive Zusammenhänge zwischen allen Prädiktoren und intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugung der Schülerinnen und Schüler an. Weiter erwarten wir negative Assoziationen aller Prädiktoren mit der Kostenüberzeugung. Basierend auf den Befunden von Hulleman und Harackiewicz (2009) gehen wir von größeren Assoziationen der wahrgenommenen Praxisorientierung im Mathematikunterricht mit der Nützlichkeitsüberzeugung aus. Da die vorliegende Studie Faktoren im natürlich auftretenden Unterrichtskontext untersucht, ist zudem generell mit schwächeren Effekten zu rechnen als in Interventionsstudien zu Wertüberzeugungen (z. B. Hulleman und Harackiewicz 2009). Darüber hinaus erwarten wir, dass die aus Lehrersicht berichteten Unterrichtsstrategien eine positive aber geringere Vorhersagekraft für die Wertüberzeugungen aufweisen als die aus Schülersicht beurteilte Praxisorientierung im Mathematikunterricht (vgl. z. B. Clausen 2002). In Bezug auf die Vorhersagekraft der Unterrichtsmerkmale für die Entwicklung der Wertüberzeugungen im Verlauf von sechs Monaten nehmen wir an, dass nur eine geringe Verstärkung der Effekte zu beobachten sein wird, da Schülerinnen und Schüler in höheren Jahrgangsstufen des Gymnasiums ihre Lehrkräfte und deren Unterrichtsstil häufig schon aus früheren Schuljahren kennen, was deren potentiellen Einfluss auf die Schülermotivation mindert.

## Methode

### Stichprobe

Zur Untersuchung der Forschungsfrage wurden Fragebogendaten aus der Interventionsstudie „Motivationsförderung im Mathematikunterricht“ (MoMa) verwendet, die im Laufe des Schuljahres 2012/13 in 82 Klassen der 9. Jahrgangsstufe an 25 Gymnasien in Baden-Württemberg erhoben wurden. Die Studienteilnahmequote lag bei 96,0 %. 1978 Schülerinnen und Schüler (53,3 % weiblich) mit aktivem Elterneinverständnis füllten zirka einen Monat (T1), zweieinhalb Monate (T2) und sieben Monate (T3) nach Schuljahresbeginn Fragebögen aus. 110 Schülerinnen und Schüler waren zum ersten Testzeitpunkt abwesend, so dass für die vorliegende Studie Daten von 1868 Schülerinnen und Schüler verwendet wurden. 72 Mathematiklehrkräfte (45,8 % weiblich) stellten zu T1 Daten zu ihren Unterrichtspraktiken zur Verfügung. Neun der Lehrkräfte unterrichteten zwei der teilnehmenden Klassen und beurteilten den Einsatz ihrer Unterrichtsstrategien getrennt für beide Klassen. Eine Lehrkraft beteiligte sich nicht an der Umfrage. Mehr als die Hälfte der Klassen (53,7 %) kannte ihre Mathematiklehrkraft bereits aus vorangegangenen Schuljahren. Die Erhebungen wurden von wissenschaftlichen Mitarbeiterinnen durchgeführt und fanden im regulären Unterricht statt.

### Instrumente

**Wertüberzeugungen.** Die Wertüberzeugungen der Schülerinnen und Schüler wurden mit vierstufigen Ratingskalen von 1 (*stimmt gar nicht*) bis 4 (*stimmt genau*) erfasst. Eingesetzt wurden sowohl etablierte Items aus früheren Studien (z. B. Steinmayr und Spinath 2010; Conley 2012; Trautwein et al. 2012) als auch neu entwickelte Items zur differenzierteren Erfassung der Wertüberzeugungen (Gaspard et al. 2015). Intrinsischer Wert wurde durch vier Items erfasst ($\alpha = .93$), Wichtigkeit durch zehn Items ($\alpha = .91$), Nützlichkeit durch 14 Items ($\alpha = .88$) und Kosten durch elf Items ($\alpha = .93$). Die Abgrenzung der vier Wertkomponenten der Schülerinnen und Schüler in Mathematik voneinander wurde durch konfirmatorische Faktorenanalysen überprüft und bestätigt (vgl. Gaspard et al. 2015).

**Unterrichtsmerkmale.** Zur Erfassung der Merkmale des Mathematikunterrichts aus Schülersicht wurden etablierte Messinstrumente mit vierstufigen Skalen von 1 (*stimmt gar nicht*) bis 4 (*stimmt genau*) verwendet (Baumert et al. 2009). Praxisorientierung wurde mit fünf Items gemessen ($\alpha = .84$) und die Wertschätzung der Mathematik im Klassenverband mit vier Items ($\alpha = .75$). Die Unterrichtsstrategien aus Lehrersicht wurden anhand von etablierten Messinstrumenten mit vierstufigen Skalen von 1 (*selten oder nie* bzw. *trifft nicht zu*) bis 4 (*regelmäßig* bzw. *trifft voll zu*) erhoben (Baumert et al., 2009). Die Themeneinführung mit Alltagsbeispielen wurde mit vier Items erfasst ($\alpha = .55$) und die Demonstration von Sachverbindungen mit fünf Items ($\alpha = .72$).

**Kovariaten.** Zur Kontrolle der Schülerleistung wurde die Diagnose- und Vergleichsarbeit in Mathematik herangezogen—ein standardisierter Test, der den Lernstand der Schülerinnen und Schüler zu Beginn der 9. Jahrgangsstufe in baden-württembergischen Gymnasien erfasst.

### Statistisches Vorgehen

**Mehrebenenregressionsanalysen.** Zur Vorhersage der Wertüberzeugungen durch die Unterrichtsmerkmale wurden mit Mplus 7 lineare Regressionsanalysen auf zwei Ebenen[1] berechnet (Muthén und Muthén 1998-2012). Die Mehrebenenmodellierung ist hierbei eine geeignete Methode, um die Gruppierung der Schülerinnen und Schüler in Klassen und die daraus resultierende hierarchische Struktur der Daten zu berücksichtigen (Raudenbush und Bryk 2002). Da durch die Variablen aus Schülersicht kein individuelles Schülerurteil sondern die in der Klasse geteilte Wahrnehmung des Unterrichts und des Wertklimas erfasst werden sollte, stehen somit die Effekte auf Klassenebene bzw. sogenannte Klimaeffekte—Unterschiede zwischen Klassen—im Fokus der Analysen (vgl. Marsh et al. 2012). Für jede Wertkomponente als abhängige Variable wurden jeweils drei Modelle berechnet. Das Modell ohne Prädiktoren stellte den ersten Analyseschritt dar (Modell 0). Danach wurden jeweils zwei multiple Regressionsanalysen modelliert. Beim Individualmodell wurden alle Prädiktoren aus Schülersicht am Klassenmittel zentriert (vgl. Enders und Tofighi 2007) auf der ersten Ebene, der Schülerebene (L1), ins Modell gefügt (Modell 1). Beim Gesamtmodell wurden die Prädiktoren aus Schülersicht zusätzlich zur Schülerebene zusammen mit den Prädiktoren aus Lehrersicht als latente Klassenaggregate (vgl. Lüdtke et al. 2008) auf der zweiten Ebene, der Klassenebene (L2), ins Modell aufgenommen (Modell 2). Die latente Aggregation wurde gewählt, um der Unreliabilität von Klassenmittelwerten Rechnung zu tragen (vgl. Marsh et al. 2012). Bei allen Regressionsmodellen wurde für das Schülergeschlecht und die Schülerleistung kontrolliert, wobei diese Kovariaten auf der Schülerebene am Klassenmittel zentriert und auf der Klassenebene als latente Aggregate in die Modelle gefügt wurden. Alle Regressionskoeffizienten wurden unter Verwendung der Gesamtvarianz der jeweiligen abhängigen Variablen und der ebenenspezifischen Varianzen der Prädiktoren standardisiert (vgl. Marsh et al. 2009). Zur Bestimmung des durch die Prädiktoren aufgeklärten Varianzanteils ($R^2$) wurde neben den in Mplus 7 integrierten Effektstärkemaßen zusätzlich das $R^2$ nach Snijders und Bosker (1994) errechnet. Dieser Determinationskoeffizient berücksichtigt die Varianz der abhängigen Variablen im Modell ohne Prädiktoren und berechnet die Varianzreduktion durch die Prädiktoren auf beiden Ebenen (Snijders und Bosker 1994).

Für die univariaten Zusammenhänge wurden für jede Wertkomponente als abhängige Variable Regressionsmodelle mit jeweils nur einem Unterrichtsmerkmal als Prädiktor auf der jeweils relevanten Ebene berechnet, wobei jeder Prädiktor aus Schülersicht simultan auf Schüler- und Klassenebene ins Modell genommen wurde. Es wurde für Geschlecht und Mathematikleistung kontrolliert. Aufgrund der Mehrfachtestung wurden die *p*-Werte der univariaten Analysen mittels Korrektur von Benjamini und Hochberg (1995) adjustiert.

Zur Überprüfung der Veränderung der Effekte im Längsschnitt wurden alle Modelle mit den Wertkomponenten zu T3 als abhängige Variablen und den Unterrichtsmerkmalen als Prädiktoren repliziert.[2] Bei diesen univariaten und multiplen Regressionsanalysen wurde für Geschlecht, Mathematikleistung und den Ausgangswert der Wertkomponenten zu T1 kontrolliert.

**Fehlende Werte.** Bei den Wertüberzeugungen und den Prädiktoren aus Schülersicht betrugen die fehlenden Werte bis zu 7,5 % zu T1 und bis zu 11,6 % zu T3. Bei Lehrkräften fehlten bis zu 2,4 % der Angaben. Entsprechend den Empfehlungen von Lüdtke et al. (2007) nutzten wir das in Mplus 7 integrierte Full Information Maximum Likelihood-Verfahren, das zur Schätzung fehlender Parameter alle zur Verfügung stehenden Informationen berücksichtigt (Enders 2010).

## Ergebnisse

### Deskriptive Statistiken

Die Mittelwerte, Standardabweichungen und Intraklassenkorrelationskoeffizienten (*ICC*) sind Tabelle 1 zu entnehmen. Wie anhand des *ICC(1)* zu erkennen ist, bestand bei den Wertüberzeugungen zu T1 bis zu 7 % und zu T2 bis zu 8 % Varianz auf Klassenebene. Außerdem variierte die Praxisorientierung im Mathematikunterricht weniger (6 %) über Klassen hinweg als die wahrgenommene Mitschülereinstellung zur Mathematik (21 %). Darüber hinaus zeigt der *ICC(2)*, dass die Reliabilität der Klassenmittelwerte der Praxisorientierung unterhalb der Grenze von .70 blieb (vgl. Lüdtke et al. 2006). In Tabelle 2 sind die latenten Interkorrelationen für die Prädiktorvariablen und die vier Wertkomponenten getrennt für die Schüler- und die Klassenebene berichtet. Es ist bemerkenswert, dass die von der Klasse im Mittel wahrgenommene Praxisorientierung im Mathematikunterricht nicht signifikant mit der von der Lehrkraft berichteten Themeneinführung mit Alltagsbeispielen korrelierte ($r$ = .21; adjustierter $p$-Wert = .208). Aufgrund der großen Varianz der wahrgenommenen Praxisorientierung im Mathematikunterricht zwischen Schülerinnen und Schülern innerhalb einer Klasse gehen wir im Folgenden auch näher auf die Individualebene ein (vgl. Lüdtke et al. 2006).

### Assoziation zwischen Unterrichtsgeschehen und Wertüberzeugungen

**Vergleich der Prädiktoren auf Schülerebene.** Unter Kontrolle von Geschlecht, Mathematikleistung und wahrgenommener Wertschätzung der Mathematik im Klassenverband zeigte sich, dass die Praxisorientierung im Mathematikunterricht auf der Schülerebene statistisch signifikant positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und negativ mit der Kostenüberzeugung zusammenhing (vgl. Tabelle 3 bis 6). Unter Kontrolle der jeweiligen Wertkomponente zu T1 wird außerdem deutlich, dass die Praxisorientierung innerhalb von sechs Monaten zu einer statistisch signifikanten Zunahme von intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und zur Abnahme von Kostenüberzeugung führte. Die Vorhersagekraft der Praxisorientierung war sowohl im Quer- als auch im Längsschnitt für die Nützlichkeitsüberzeugung am stärksten und für die Kostenüberzeugung am geringsten. Die wahrgenommene Wertschätzung der Mathematik im Klassenverband war unter Kontrolle der anderen Variablen auf Schülerebene statistisch signifikant positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und negativ mit der Kostenüberzeugung assoziiert. Der Zusammenhang war mit intrinsischem Wert am engsten und mit der Kostenüberzeugung am niedrigsten. Interessanterweise trug die wahrgenommene Wertschätzung der Mathematik im Klassenverband im Längsschnitt zu einer statistisch signifikanten Zunahme der Nützlichkeitsüberzeugung bei. Gemessen an den Effektstärken spielte die Praxisorientierung im Quer- und im Längsschnitt für alle Wertkomponenten eine stärkere Rolle als die wahrgenommene Wertschätzung der Mathematik im Klassenverband.

Die Prädiktoren und Kovariaten auf der Schülerebene trugen zur meisten Varianzaufklärung bei der Nützlichkeitsüberzeugung bei. Die Regressions- und Determinationskoeffizienten in den Individualmodellen und den Gesamtmodellen unterschieden sich insgesamt kaum.

**Vergleich der Prädiktoren auf Klassenebene.** Die im Klassenmittel wahrgenommene Praxisorientierung war unter Kontrolle von Geschlecht, Mathematikleistung und den anderen Prädiktoren auf Klassenebene statistisch signifikant positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und marginal signifikant negativ mit der Kostenüberzeugung assoziiert, trug jedoch nicht zu statistisch signifikanten Veränderungen der Wertüberzeugungen nach sechs Monaten bei (vgl. Tabelle 3 bis 6). Der Zusammenhang der mittleren Praxisorientierung war mit der Nützlichkeitsüberzeugung deutlich enger als mit den anderen Wertkomponenten. Die im Klassenmittel wahrgenommene Wertschätzung der Mathematik im Klassenverband hing unter Kontrolle der anderen Variablen auf Klassenebene statistisch signifikant positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und negativ mit der Kostenüberzeugung zusammen, wobei sich der Zusammenhang mit intrinsischem Wert als am engsten erwies. Interessanterweise waren die Zusammenhänge der mittleren wahrgenommenen Wertschätzung der Mathematik im Klassenverband im Querschnitt mit fast allen Wertkomponenten stärker als jene der Praxisorientierung. Allerdings war auch die wahrgenommene Wertschätzung der Mathematik im Klassenverband auf Klassenebene nicht mit Veränderungen der Wertüberzeugungen über sechs Monate assoziiert.

Unter Kontrolle der Kovariaten und der anderen Prädiktoren auf Klassenebene konnte die von den Lehrkräften berichtete Themeneinführung mit Alltagsbeispielen im Querschnitt zwar keine Varianz in den Wertüberzeugungen der Schülerinnen und Schüler aufklären, führte jedoch zu einer statistisch signifikanten Abnahme der Kostenüberzeugung nach sechs Monaten (vgl. Tabelle 3 bis 6). Die Demonstration von Sachverbindungen war im Querschnitt unter Kontrolle der anderen Variablen statistisch signifikant positiv mit der Wichtigkeitsüberzeugung und marginal signifikant negativ mit der Kostenüberzeugung assoziiert, nicht aber mit intrinsischem Wert und der Nützlichkeitsüberzeugung. Erwartungskonform war die Vorhersagekraft dieses Prädiktors für die Wichtigkeits- und Kostenüberzeugungen geringer als jene der gemittelten Prädiktoren aus Schülersicht. Im Längsschnitt zeigten sich keine statistisch signifikanten Effekte der Demonstration von Sachverbindungen auf die Entwicklung der Wertüberzeugungen.

Durch die Prädiktoren und Kovariaten konnten 87 % (intrinsischer Wert) bis 99 % (Kostenüberzeugung) der Varianzen auf Klassenebene aufgeklärt werden. Das $R^2$ nach Snijders und Bosker (1994) ist gerade vor dem Hintergrund der geringen Zwischenklassenvarianzen in den Wertüberzeugungen ein weiteres wichtiges Effektstärkemaß und zeigt, dass sich die Gesamtvarianz im Gesamtmodell im Vergleich zum Nullmodell bei der Nützlichkeitsüberzeugung am stärksten, nämlich um 52 %, reduzierte.

Tabelle 1
*Deskriptive Statistiken*

| Messzeitpunkt | | T1 | | | | T3 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Beispielitem | *N* | *M* (*SD*) | ICC(1) | ICC(2) | *N* | *M* (*SD*) | ICC(1) | ICC(2) |
| *Wertüberzeugungen in Mathematik* | | | | | | | | | |
| Intrinsischer Wert | Mathematik macht mir Spaß. | 1849 | 2,26 (0,85) | 0,07 | 0,44 | 1651 | 2,34 (0,81) | 0,08 | 0,44 |
| Wichtigkeitsüberzeugung | Es ist mir wichtig, gut in Mathe zu sein. | 1858 | 2,78 (0,59) | 0,04 | 0,33 | 1657 | 2,83 (0,61) | 0,06 | 0,38 |
| Nützlichkeitsüberzeugung | Mathematik ist sehr nützlich für mich. | 1860 | 2,52 (0,49) | 0,06 | 0,43 | 1661 | 2,51 (0,51) | 0,08 | 0,46 |
| Kostenüberzeugung | Ich muss viel aufgeben, um in Mathe gut zu sein. | 1859 | 2,11 (0,69) | 0,04 | 0,28 | 1660 | 2,09 (0,73) | 0,06 | 0,39 |
| *Merkmale des Mathematikunterrichts aus Schülersicht* | | | | | | | | | |
| Praxisorientierung | Im Mathematikunterricht beschäftigen wir uns mit Aufgaben, die einen praktischen Nutzen haben. | 1847 | 2,43 (0,63) | 0,06 | 0,42 | – | – – | – | – |
| Wertschätzung des Fachs Mathematik im Klassenverband | Die meisten Schüler/innen in meiner Klasse halten das Fach Mathematik für wichtig. | 1829 | 1,93 (0,55) | 0,21 | 0,73 | – | – – | – | – |
| *Merkmale des Mathematikunterrichts aus Lehrersicht* | | | | | | | | | |
| Themeneinführung mit Alltagsbeispielen | Wenn ich einen neuen Begriff oder Sachverhalt einführe, gehe ich von einem Alltagsproblem aus, bei dem sich die Notwendigkeit ergibt, den neuen Begriff einzuführen und zu definieren. | 80 | 2,67 (0,52) | n. z. | n. z. | – | – – | – | – |
| Demonstration von Sachverbindungen | Ich bringe den Mathematikstoff mit Dingen in Verbindung, die die Schüler/innen in anderen Fächern gelernt haben. | 81 | 3,04 (0,48) | n. z. | n. z. | – | – – | – | – |

Der ICC(1) gibt den prozentualen Anteil der Gesamtvarianz an, der auf Unterschiede zwischen den Klassen zurückzuführen ist (Lüdtke et al. 2006, S. 87). Der ICC(2) gibt die Reliabilität des Klassenmittelwerts der Schülerurteile an (Lüdtke et al. 2006, S. 87).
*N* = Anzahl; *M* = Mittelwert, *SD* = Standardabweichung, ICC = Intraklassenkorrelationskoeffizient, n. z. = nicht zutreffend.

Tabelle 2

*Latente Interkorrelationen auf Schülerebene (rechts oberhalb der Diagonalen) und auf Klassenebene (links unterhalb der Diagonalen)*

| Messzeitpunkt | T1 | | | | | | | T3 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 Intrinsischer Wert | – | 0,61*** | 0,54*** | -0,67*** | 0,48*** | 0,26*** | – | – | 0,57*** | 0,50*** | -0,56*** | 0,40*** | 0,21*** | – |
| 2 Wichtigkeitsüberzeugung | 0,92*** | – | 0,67*** | -0,38*** | 0,48*** | 0,22*** | – | 0,73*** | – | 0,65*** | -0,34*** | 0,39*** | 0,18*** | – |
| 3 Nützlichkeitsüberzeugung | 0,83*** | 0,95*** | – | -0,31*** | 0,68*** | 0,29*** | – | 0,78*** | 0,93*** | – | -0,23*** | 0,47*** | 0,23*** | – |
| 4 Kostenüberzeugung | -0,93*** | -0,89*** | -0,83*** | – | -0,37*** | -0,12*** | – | -0,73*** | -0,31† | -0,40* | – | -0,32*** | -0,09*** | – |
| 5 Praxisorientierung (S) | 0,81*** | 0,86*** | 0,93*** | -0,72*** | – | 0,26*** | – | 0,80*** | 0,76*** | 0,85*** | -0,54*** | – | 0,26*** | – |
| 6 Wertschätzung des Fachs im Klassenverband (S) | 0,93*** | 0,92*** | 0,85*** | -0,97*** | 0,76*** | – | – | 0,86*** | 0,74*** | 0,71*** | -0,75*** | 0,77*** | – | – |
| 7 Themeneinführung mit Alltagsbeispielen (L) | 0,08 | 0,17 | 0,17 | -0,02 | 0,21 | 0,15 | – | 0,20 | 0,23† | 0,17 | -0,32† | 0,21 | 0,15 | – |
| 8 Demonstration von Sachverbindungen (L) | 0,24 | 0,37** | 0,33* | -0,01 | 0,36* | 0,04 | 0,23* | 0,34** | 0,24† | 0,33** | -0,01 | 0,35* | 0,04 | 0,25* |

Skalen auf Schülerebene wurden am Klassenmittel zentriert. Aufgrund der Mehrfachtestung wurden die *p*-Werte adjustiert (Korrektur von Benjamini und Hochberg, 1995).
Alle Variablen beziehen sich auf das Fach Mathematik. S = Schülerangabe, L = Lehrerangabe.
*** $p < 0{,}002$; ** $p < 0{,}01$; * $p < 0{,}05$; † $p < 0{,}10$.

Tabelle 3

*Vorhersage von intrinsischem Wert in Mathematik durch relevanzbezogene Merkmale des Mathematikunterrichts*

| Messzeitpunkt | T1 | | | | | | T3 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Modell | 0 | | 1 | | 2 | | 0 | | 1 | | 2 | |
| | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) |
| L1[a] Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | 0,07 | (0,02) | 0,07 | (0,02) | | | 0,08 *** | (0,02) | 0,08 *** | (0,02) |
| Mathematikleistung | | | 0,33 *** | (0,02) | 0,33 *** | (0,03) | | | 0,05 ** | (0,02) | 0,05 ** | (0,02) |
| Intrinsischer Wert T1 | | | n. z. | | n. z. | | | | 0,65 *** | (0,03) | 0,63 *** | (0,02) |
| L1[a] Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung | | | 0,37 *** | (0,03) | 0,36 *** | (0,02) | | | 0,08 *** | (0,02) | 0,08 *** | (0,02) |
| Wertschätzung Klassenverband (S) | | | 0,16 *** | (0,02) | 0,16 *** | (0,02) | | | 0,02 | (0,02) | 0,02 | (0,02) |
| L2 Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | | | 0,03 | (0,02) | | | | | 0,02 | (0,03) |
| Mathematikleistung | | | | | 0,02 | (0,02) | | | | | -0,01 | (0,03) |
| Intrinsischer Wert T1 | | | | | n. z. | | | | | | 0,15 ** | (0,06) |
| L2 Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung (S) | | | | | 0,08 * | (0,03) | | | | | 0,02 | (0,04) |
| Wertschätzung Klassenverband (S) | | | | | 0,16 *** | (0,03) | | | | | 0,06 | (0,04) |
| Alltagsbeispiele (L) | | | | | -0,03 | (0,02) | | | | | 0,02 | (0,02) |
| Sachverbindungen (L) | | | | | 0,02 | (0,02) | | | | | 0,04 | (0,02) |
| Varianzkomponenten | | | | | | | | | | | | |
| Schülerebene | 0,672 *** | (0,02) | 0,423 *** | (0,01) | 0,423 *** | (0,01) | 0,601 *** | (0,02) | 0,269 *** | (0,01) | 0,269 *** | (0,01) |
| Klassenebene | 0,048 *** | (0,01) | | | 0,004 | (0,00) | 0,048 *** | (0,01) | | | 0,011 ** | (0,00) |
| $R^2$ auf Schülerebene | | | 0,319 *** | (0,02) | 0,312 *** | (0,02) | | | 0,512 *** | (0,02) | 0,499 *** | (0,02) |
| $R^2$ auf Klassenebene | | | | | 0,870 *** | (0,12) | | | | | 0,648 *** | (0,11) |
| Snijders & Bosker (1994) $R^2_{L1}$ | | | | | 0,407 | | | | | | 0,569 | |

[a] Skalen wurden am Klassenmittel zentriert. L1 = Level 1 (Schülerebene), L2 = Level 2 (Klassenebene), m = männlich; S = Schülerangabe, L = Lehrerangabe, $\beta$ = standardisierter Regressionskoeffizient, SE = Standardfehler, $R^2$ = aufgeklärte Varianz, n. z. = nicht zutreffend.
*** $p < 0{,}001$; ** $p < 0{,}01$; * $p < 0{,}05$; † $p < 0{,}10$.

Tabelle 4

*Vorhersage von Wichtigkeitsüberzeugung in Mathematik durch relevanzbezogene Merkmale des Mathematikunterrichts*

| Messzeitpunkt | T1 | | | | | | T3 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Modell | 0 | | 1 | | 2 | | 0 | | 1 | | 2 | |
| | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) |
| L1[a] Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | 0,02 | (0,02) | 0,02 | (0,02) | | | -0,03 * | (0,02) | -0,03 † | (0,02) |
| Mathematikleistung | | | 0,17 *** | (0,03) | 0,17 *** | (0,03) | | | 0,06 ** | (0,02) | 0,07 ** | (0,02) |
| Wichtigkeitsüberzeugung T1 | | | n. z. | | n. z. | | | | 0,61 *** | (0,02) | 0,60 *** | (0,03) |
| L1[a] Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung | | | 0,42 *** | (0,03) | 0,41 *** | (0,02) | | | 0,08 *** | (0,02) | 0,08 *** | (0,02) |
| Wertschätzung Klassenverband (S) | | | 0,11 *** | (0,02) | 0,11 *** | (0,02) | | | 0,03 | (0,03) | 0,03 | (0,02) |
| L2 Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | | | -0,03 | (0,02) | | | | | -0,06 * | (0,03) |
| Mathematikleistung | | | | | -0,01 | (0,02) | | | | | -0,02 | (0,03) |
| Wichtigkeitsüberzeugung T1 | | | | | n. z. | | | | | | 0,12 † | (0,07) |
| L2 Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung (S) | | | | | 0,08 * | (0,03) | | | | | 0,07 | (0,05) |
| Wertschätzung Klassenverband (S) | | | | | 0,13 *** | (0,03) | | | | | 0,02 | (0,04) |
| Alltagsbeispiele (L) | | | | | -0,01 | (0,02) | | | | | 0,02 | (0,02) |
| Sachverbindungen (L) | | | | | 0,05 ** | (0,02) | | | | | 0,01 | (0,03) |
| Varianzkomponenten | | | | | | | | | | | | |
| Schülerebene | 0,332 *** | (0,01) | 0,242 *** | (0,01) | 0,241 *** | (0,01) | 0,350 *** | (0,01) | 0,185 *** | (0,01) | 0,185 *** | (0,01) |
| Klassenebene | 0,015 ** | (0,01) | | | 0,000 | (0,00) | 0,022 *** | (0,01) | | | 0,005 * | (0,00) |
| $R^2$ auf Schülerebene | | | 0,240 *** | (0,02) | 0,231 *** | (0,02) | | | 0,437 *** | (0,03) | 0,425 *** | (0,03) |
| $R^2$ auf Klassenebene | | | | | 0,953 *** | (0,20) | | | | | 0,683 *** | (0,17) |
| Snijders & Bosker (1994) $R^2_{L1}$ | | | | | 0,305 | | | | | | 0,489 | |

[a] Skalen wurden am Klassenmittel zentriert. L1 = Level 1 (Schülerebene), L2 = Level 2 (Klassenebene), m = männlich; S = Schülerangabe, L = Lehrerangabe, $\beta$ = standardisierter Regressionskoeffizient, SE = Standardfehler, $R^2$ = aufgeklärte Varianz, n. z. = nicht zutreffend.

*** $p < 0{,}001$; ** $p < 0{,}01$; * $p < 0{,}05$; † $p < 0{,}10$.

Tabelle 5

*Vorhersage von Nützlichkeitsüberzeugung in Mathematik durch relevanzbezogene Merkmale des Mathematikunterrichts*

| Messzeitpunkt | T1 | | | | | | T3 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Modell | 0 | | 1 | | 2 | | 0 | | 1 | | 2 | |
| | β | (SE) | β | (SE) | β | (SE) | β | (SE) | β | (SE) | β | (SE) |
| L1 [a] Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | 0,01 | (0,02) | 0,01 | (0,02) | | | 0,03 † | (0,02) | 0,03 | (0,02) |
| Mathematikleistung | | | 0,11 *** | (0,02) | 0,11 *** | (0,02) | | | 0,05 ** | (0,02) | 0,05 ** | (0,02) |
| Nützlichkeitsüberzeugung T1 | | | n. z. | | n. z. | | | | 0,52 *** | (0,03) | 0,50 *** | (0,03) |
| L1 [a] Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung | | | 0,63 *** | (0,02) | 0,61 *** | (0,02) | | | 0,09 *** | (0,03) | 0,09 *** | (0,02) |
| Wertschätzung Klassenverband (S) | | | 0,13 *** | (0,02) | 0,12 *** | (0,02) | | | 0,05 * | (0,02) | 0,05 * | (0,02) |
| L2 Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | | | -0,01 | (0,02) | | | | | 0,01 | (0,03) |
| Mathematikleistung | | | | | -0,01 | (0,02) | | | | | -0,04 | (0,03) |
| Nützlichkeitsüberzeugung T1 | | | | | n. z. | | | | | | 0,18 ** | (0,06) |
| L2 Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung (S) | | | | | 0,16 *** | (0,04) | | | | | 0,06 | (0,06) |
| Wertschätzung Klassenverband (S) | | | | | 0,10 ** | (0,03) | | | | | 0,01 | (0,04) |
| Alltagsbeispiele (L) | | | | | -0,01 | (0,02) | | | | | 0,01 | (0,02) |
| Sachverbindungen (L) | | | | | 0,03 | (0,02) | | | | | 0,02 | (0,03) |
| Varianzkomponenten | | | | | | | | | | | | |
| Schülerebene | 0,228 *** | (0,01) | 0,116 *** | (0,00) | 0,116 *** | (0,00) | 0, 238 *** | (0,01) | 0,145 *** | (0,01) | 0,145 *** | (0,01) |
| Klassenebene | 0,016 *** | (0,00) | | | 0,001 | (0,00) | 0,021 *** | (0,01) | | | 0,003 * | (0,00) |
| $R^2$ auf Schülerebene | | | 0,471 *** | (0,02) | 0,453 *** | (0,02) | | | 0,338 *** | (0,03) | 0,323 *** | (0,03) |
| $R^2$ auf Klassenebene | | | | | 0,928 *** | (0,08) | | | | | 0,811 *** | (0,10) |
| Snijders & Bosker (1994) $R^2_{L1}$ | | | | | 0,520 | | | | | | 0,429 | |

[a] Skalen wurden am Klassenmittel zentriert. L1 = Level 1 (Schülerebene), L2 = Level 2 (Klassenebene), m = männlich; S = Schülerangabe, L = Lehrerangabe, β = standardisierter Regressionskoeffizient, SE = Standardfehler, $R^2$ = aufgeklärte Varianz, n. z. = nicht zutreffend.

*** $p < 0{,}001$; ** $p < 0{,}01$; * $p < 0{,}05$; † $p < 0{,}10$.

Tabelle 6

*Vorhersage von Kostenüberzeugung in Mathematik durch relevanzbezogene Merkmale des Mathematikunterrichts*

| Messzeitpunkt | T1 | | | | | | T3 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Modell | 0 | | 1 | | 2 | | 0 | | 1 | | 2 | |
| | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) |
| L1 [a] Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | -0,06 ** | (0,02) | -0,06 ** | (0,02) | | | -0,04 * | (0,02) | -0,04 * | (0,02) |
| Mathematikleistung | | | -0,39 *** | (0,02) | -0,39 *** | (0,02) | | | -0,11 ** | (0,02) | -0,11 *** | (0,02) |
| Kostenüberzeugung T1 | | | n. z. | | n. z. | | | | 0,66 *** | (0,02) | 0,64 *** | (0,03) |
| L1 [a] Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung | | | -0,28 *** | (0,02) | -0,27 *** | (0,02) | | | -0,07 *** | (0,02) | -0,07 *** | (0,02) |
| Wertschätzung Klassenverband (S) | | | -0,05 * | (0,02) | -0,04 * | (0,02) | | | 0,01 | (0,02) | 0,01 | (0,02) |
| L2 Kontrollvariablen | | | | | | | | | | | | |
| Geschlecht (1 = m) | | | | | -0,02 | (0,02) | | | | | -0,01 | (0,02) |
| Mathematikleistung | | | | | -0,05 | (0,03) | | | | | -0,02 | (0,03) |
| Kostenüberzeugung T1 | | | | | n. z. | | | | | | 0,20 ** | (0,06) |
| L2 Unterrichtsmerkmale | | | | | | | | | | | | |
| Praxisorientierung (S) | | | | | -0,06 † | (0,03) | | | | | 0,01 | (0,04) |
| Wertschätzung Klassenverband (S) | | | | | -0,11 *** | (0,03) | | | | | -0,01 | (0,04) |
| Alltagsbeispiele (L) | | | | | 0,02 | (0,02) | | | | | -0,07 ** | (0,02) |
| Sachverbindungen (L) | | | | | 0,03 † | (0,02) | | | | | 0,02 | (0,03) |
| Varianzkomponenten | | | | | | | | | | | | |
| Schülerebene | 0,462 *** | (0,01) | 0,328 *** | (0,01) | 0,328 *** | (0,01) | 0, 238 *** | (0,01) | 0,226 *** | (0,01) | 0,226 *** | (0,01) |
| Klassenebene | 0,016 ** | (0,01) | | | 0,000 | (0,00) | 0,021 *** | (0,01) | | | 0,005 | (0,00) |
| $R^2$ auf Schülerebene | | | 0,257 *** | (0,02) | 0,250 *** | (0,01) | | | 0,515 *** | (0,02) | 0,502 *** | (0,03) |
| $R^2$ auf Klassenebene | | | | | 0,994 ** | (0,38) | | | | | 0,821 *** | (0,16) |
| Snijders & Bosker (1994) $R^2_{L1}$ | | | | | 0,314 | | | | | | 0,570 | |

[a] Skalen wurden am Klassenmittel zentriert. L1 = Level 1 (Schülerebene), L2 = Level 2 (Klassenebene), m = männlich; S = Schülerangabe, L = Lehrerangabe, $\beta$ = standardisierter Regressionskoeffizient, SE = Standardfehler, $R^2$ = aufgeklärte Varianz, n. z. = nicht zutreffend.
*** $p < 0{,}001$; ** $p < 0{,}01$; * $p < 0{,}05$; † $p < 0{,}10$.

Tabelle 7

*Univariate Zusammenhänge relevanzbezogener Unterrichtsmerkmale mit Wertüberzeugungen in Mathematik zu T1 und T3 (unter Kontrolle der Wertüberzeugung zu T1)*

| | Intrinsischer Wert | | | | Wichtigkeitsüberzeugung | | | | Nützlichkeitsüberzeugung | | | | Kostenüberzeugung | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Messzeitpunkt | T1 | | T3 | | T1 | | T3 | | T1 | | T3 | | T1 | | T3 | |
| Variable | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) | $\beta$ | (SE) |
| L1: Schülerebene | | | | | | | | | | | | | | | | |
| Praxisorientierung [a] (S) | 0,40*** | (0,02) | 0,08*** | (0,02) | 0,44*** | (0,02) | 0,09*** | (0,02) | 0,64*** | (0,02) | 0,10*** | (0,02) | -0,29*** | (0,02) | -0,06*** | (0,02) |
| Wertschätzung im Klassenverband [a] (S) | 0,25*** | (0,02) | 0,03 | (0,02) | 0,21*** | (0,02) | 0,04$^{\dagger}$ | (0,02) | 0,28*** | (0,02) | 0,06* | (0,02) | -0,11*** | (0,02) | 0,00 | (0,01) |
| L2: Klassenebene | | | | | | | | | | | | | | | | |
| Praxisorientierung [a] (S) | 0,19*** | (0,03) | 0,05 | (0,04) | 0,18*** | (0,04) | 0,08 | (0,03) | 0,24*** | (0,03) | 0,06 | (0,06) | -0,12*** | (0,03) | 0,00 | (0,04) |
| Wertschätzung im Klassenverband [a] (S) | 0,22*** | (0,03) | 0,05 | (0,04) | 0,20*** | (0,03) | 0,04 | (0,04) | 0,22*** | (0,03) | 0,01 | (0,04) | -0,16*** | (0,03) | -0,02 | (0,03) |
| Alltagsbeispiele (L) | 0,02 | (0,03) | 0,03 | (0,02) | 0,04 | (0,02) | 0,03 | (0,02) | 0,04 | (0,03) | 0,02 | (0,02) | 0,00 | (0,03) | -0,06* | (0,02) |
| Sachverbindungen (L) | 0,03 | (0,03) | 0,04$^{\dagger}$ | (0,02) | 0,07* | (0,03) | 0,01 | (0,02) | 0,07* | (0,03) | 0,02 | (0,03) | 0,03 | (0,02) | 0,00 | (0,02) |

Aufgrund der Mehrfachtestung wurden die *p*-Werte adjustiert (Korrektur von Benjamini und Hochberg, 1995).

[a] Prädiktor wurde simultan auf Schülerebene (Zentrierung am Klassenmittel, vgl. Enders und Tofighi 2007) und als latentes Aggregat auf Klassenebene (vgl. Lüdtke et al. 2008) ins Modell eingefügt. L1 = Level 1 (Schülerebene), L2 = Level 2 (Klassenebene), S = Schülerangabe, L = Lehrerangabe, $\beta$ = standardisierter Regressionskoeffizient, SE = Standardfehler.

*** $p < 0{,}002$; * $p < 0{,}05$; $^{\dagger}$ $p < 0{,}10$.

**Vorhersagekraft der Prädiktoren im Einzelmodell.** Bei den Einzelassoziationen der Unterrichtsmerkmale mit den Wertkomponenten (Tabelle 7) zeigte sich für die Praxisorientierung im Mathematikunterricht, dass die einfachen Regressionskoeffizienten auf Klassenebene im Querschnitt höher ausfielen als die partiellen Regressionskoeffizienten nach Kontrolle der anderen Variablen. Bei der wahrgenommenen Wertschätzung der Mathematik im Klassenverband war dies auf beiden Ebenen der Fall. Weiter zeigte sich, dass die Themeneinführung mit Alltagsbeispielen auch einzeln im Querschnitt zwar mit keiner der Wertüberzeugungen assoziiert war, jedoch eine statistisch signifikante Abnahme der Kostenüberzeugung nach sechs Monaten vorhersagte. Ungleich den Ergebnissen in den Gesamtmodellen hing die Demonstration von Sachverbindungen im Querschnitt statistisch signifikant positiv mit der Wichtigkeitsüberzeugung und der Nützlichkeitsüberzeugung zusammen, nicht aber mit den anderen Wertkomponenten, und sagte im Längsschnitt eine marginal signifikante Zunahme von intrinsischem Wert vorher.

## Diskussion

Etliche Mathematiklehrkräfte der Sekundarstufe beschäftigt die Frage, welche Unterrichtsfaktoren zu einer positiven Einstellung ihrer Schülerinnen und Schüler zur Mathematik beitragen (vgl. Pierce und Stacey 2006). Anhand eines Datensatzes mit baden-württembergischen Gymnasiastinnen und Gymnasiasten der neunten Jahrgangsstufe wurden in der vorliegenden Studie die Relevanzorientierung im Mathematikunterricht und die wahrgenommene Wertschätzung der Mathematik im Klassenverband als Prädiktoren der Wertüberzeugungen von Schülerinnen und Schülern untersucht.

### Lehrer- und Mitschülerverhalten und Wertüberzeugungen von Schülerinnen und Schüler

Die Ergebnisse unserer Studie zeigen, dass die aus Schülersicht beurteilte Praxisorientierung und die wahrgenommene Wertschätzung des Fachs Mathematik im Klassenverband positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugungen und negativ mit der Kostenüberzeugung assoziiert sind und zu Veränderungen in den Wertüberzeugungen innerhalb von sechs Monaten auf Schülerebene führen, jedoch nicht auf Klassenebene. Die aus Lehrersicht berichtete Demonstration von Sachverbindungen war positiv mit der Wichtigkeitsüberzeugung assoziiert und tendierte zu einem positiven Zusammenhang mit der Kostenüberzeugung, sagte aber nicht die Entwicklung der klassenspezifischen Wertüberzeugungen vorher. Die von Lehrkräften berichtete Demonstration von Alltagsbeispielen im Unterricht führte dahingegen zu einer Abnahme der in der Klasse wahrgenommenen Kostenüberzeugung. Die Ergebnisse verdeutlichen, dass relevanzbezogene Unterrichtsmerkmale nicht nur mit einzelnen (z. B. Frenzel et al. 2010; Wang 2012), sondern mit allen Wertüberzeugungen von Schülerinnen und Schülern in Mathematik zusammenhängen, wobei die Effekte je nach Wertkomponente und Analyseebene variieren.

Zwar standen die Vorhersagekraft der gemittelten Schülerurteile der Unterrichtsmerkmale für die Wertüberzeugungen und somit die Klimaeffekte im Vordergrund der Analysen

(Marsh et al. 2012), die geringe Übereinstimmung der Schülerwahrnehmungen zur Praxisorientierung im Mathematikunterricht innerhalb einer Klasse legt allerdings nahe, auch auf die Schülerebene näher einzugehen (vgl. Lüdtke et al. 2006). Dass die Praxisorientierung im Mathematikunterricht besonders wichtig für die individuellen Wertüberzeugungen der Schülerinnen und Schüler ist sowie für deren Entwicklung im Verlauf von sechs Monaten, steht im Einklang mit Befunden zur Rolle der wahrgenommenen Relevanz des Unterrichts für die individuelle Interessensentwicklung in Mathematik (Willems 2011). Zieht man die geringe Varianz zwischen Klassen bei der Beurteilung der Praxisorientierung und die niedrigen latenten Korrelationen mit den von der Lehrkraft berichteten Unterrichtsstrategien mit in Betracht, so deuten unsere Ergebnisse an, dass die Wahrnehmung der Praxisorientierung von individuellen Faktoren abhängt—beispielsweise davon, ob sich die Schülerinnen und Schüler bei der Beurteilung auf ihren persönlich erlebten Alltag oder auf eine allgemeine Anwendungsorientierung beziehen (vgl. z. B. Vorhölter 2009; Vorhölter und Vollstedt 2012). Diese Wahrnehmung scheint mitunter dauerhaft darüber zu entscheiden, wie interessant, wichtig, nützlich und anstrengend bzw. nervenaufreibend sie individuell das Mathematiklernen empfinden.

In Bezug auf die Klimaeffekte ist bemerkenswert, dass die Schülerinnen und Schüler auf Klassenebene der Mathematik insbesondere mehr Nutzen zuschreiben, je mehr Praxisorientierung sie im Unterricht wahrnehmen. Für Wichtigkeits- und Kostenüberzeugungen und vor allem für intrinsischen Wert spielt dahingegen auf Klassenebene das wahrgenommene klassenspezifische „Werteklima“ gegenüber der Mathematik eine stärkere Rolle als die inhaltliche Orientierung des Mathematikunterrichts. Die starke Vorhersagekraft dieses sozialnormativen Unterrichtsmerkmals bestätigt die zunehmende Bedeutung von Gleichaltrigen im Sekundärschulalter (vgl. z. B. Parker et al. 2006). Glauben Schülerinnen und Schüler, dass ihre Klassenkameradinnen und -kameraden Mathematik mögen und für wichtig halten, so sind sie möglicherweise deswegen eher vom Wert der Mathematik überzeugt, weil sie sich an der Haltung der anderen orientieren. Dass eine als positiv beurteilte Einstellung der Klasse zur Mathematik auch damit zusammenhängt, dass in der Klasse Mathematik als weniger anstrengend und nervenaufreibend empfunden wurde, könnte darin begründet sein, dass eine positive und unterstützende Lernatmosphäre entsteht, wenn eine hohe Wertschätzung für die Mathematik im Klassenverband vorherrscht. Dass das Wertklima im Klassenverband stark mit intrinsischem Wert und Wichtigkeitsüberzeugung in Mathematik assoziiert ist, aber dennoch nicht zu deren Veränderung führt, entspricht mehrheitlich den Ergebnissen aus der Forschung zu Interesse in Mathematik (Frenzel et al. 2010) und zu schulischen Werten im Freundeskreis und Wertüberzeugungen (Ryan 2001). Bemerkenswert ist jedoch, dass die wahrgenommene Wertschätzung der Mathematik im Klassenverband mit einer Zunahme der individuellen Nützlichkeitsüberzeugung einherging. Es scheint also, als wären die Peers zumindest auf Individualebene bedeutungsvoll dafür, wie sich die Beurteilung der Relevanz der Mathematik durch Sekundarschülerinnen und -schüler entwickelt.

Die aus Lehrerperspektive erfassten Prädiktoren wiesen, wie erwartet, schwächere Zusammenhänge mit den Wertüberzeugungen der Schülerinnen und Schüler auf als jene aus Schülersicht. Überraschend sind hierbei inkonsistente Befunde zur Themeneinführung mit Alltagsbeispielen: Im Querschnitt zeigten sich keine Zusammenhänge mit den Wertüberzeugungen der Schülerinnen und Schüler in Mathematik (vgl. im Gegensatz dazu Rakoczy et al. 2008), im Längsschnitt deutete sich jedoch die Bedeutung dieser Unterrichtsstrategie für die Abnahme von Kostenüberzeugung an. Die fehlende Assoziation im Querschnitt könnte darin begründet sein, dass das, was Lehrkräfte als alltagsnah empfinden, möglicherweise nur wenig mit der Schülerperspektive übereinstimmt (vgl. dazu z. B. Clausen 2002); so zeigte sich auch kein statistisch signifikanter Zusammenhang von wahrgenommener Praxisorientierung im Mathematikunterricht aus Schülersicht und der Themeneinführung mit Alltagsbeispielen aus Lehrersicht. Denkbar ist außerdem, dass kein linearer Zusammenhang vorliegt: So könnte eine zu starke Themeneinführung mit Alltagsbeispielen das Bedürfnis von Schülerinnen und Schülern nach einer strukturierten analytischen Herangehensweise untermauern (vgl. Pierce und Stacey 2006). Eine Balance zwischen der Verwendung von Alltagsbeispielen, um Interesse zu wecken, und Formeln, um Strukturiertheit zu schaffen, könnte möglicherweise mit höheren Wertüberzeugungen in Mathematik einhergehen (vgl. Pierce und Stacey 2006; Turner und Meyer 2009). Gleichzeitig ist die positive Vorhersagekraft der Themeneinführung mit Alltagsbeispielen für die Zunahme von Kostenüberzeugungen unter Berücksichtigung des tendenziell negativen Zusammenhangs der Demonstration von Sachverbindungen mit Kostenüberzeugungen durchaus interessant: So ist die Verwendung von Aufgaben mit Alltagsbezug in Mathematik möglicherweise ein weniger anspruchsvolles Unterrichtsvorgehen als die Herstellung von inhaltlichen Bezügen zu anderen Fächern und birgt somit weniger die Gefahr der kognitiven Überforderung von Schülerinnen und Schülern (vgl. auch Kunter und Voss 2013). Dies könnte begründen, dass die untersuchten Unterrichtsstrategien eher mit der Kostenüberzeugung als mit anderen Wertkomponenten in Verbindung stehen.

Zuletzt ist zu erwähnen, dass durch alle untersuchten Unterrichtsmerkmale zusammen im Gesamtmodell gut die Hälfte und damit am meisten Gesamtvarianz in der Nützlichkeitswahrnehmung aufgeklärt werden konnte. Dieser enge Zusammenhang mit externen Unterrichtsfaktoren bestätigt die extrinsische Natur dieser Wertkomponenten (vgl. Wigfield et al. 2009) und unterstreicht das Potential von Interventionen zur Motivationsförderung im Mathematikunterricht, die auf Nützlichkeitsüberzeugungen abzielen (vgl. Hulleman und Harackiewicz 2009).

### Stärken, Grenzen und zukünftige Forschung

Diese Studie erforschte anhand eines umfangreichen Datensatzes, unter welchen Bedingungen im Mathematikunterricht Schülerinnen und Schüler hohe Wertüberzeugungen in Mathematik haben. Eine der wesentlichen Stärken dieser Studie ist dabei, dass sowohl die Schüler- als auch die Lehrerperspektive auf das Unterrichtsgeschehen in Betracht gezogen wurden. Außerdem wurden die Wertkomponenten intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung simultan in einem elaborierten Fragebogen erfasst.

Die Ergebnisse dieser Studie können allerdings nur in Verbindung mit den spezifischen Charakteristika unserer Stichprobe interpretiert werden. So waren im Laufe von sechs Monaten kaum Effekte der untersuchten Unterrichtsmerkmale auf die motivationale Entwicklung der Schülerinnen und Schüler zu beobachten; da aber die Mehrheit der untersuchten Gymnasialklassen ihre Mathematiklehrkraft bereits aus dem Vorjahr kannten, ist zu vermuten, dass sich die Effekte des Unterrichts zu Schuljahresbeginn bereits niedergeschlagen hatten und daher Effekte auf weitere motivationale Veränderungen ausblieben.[3] Die Frage nach Kausalität lässt sich somit anhand des untersuchten Datensatzes nicht ausreichend beantworten. Weitere Längsschnittstudien sind notwendig, die die Wirkung relevanzbezogener Unterrichtsmerkmale auf die Wertüberzeugungen von Schülerinnen und Schülern in Mathematik sowohl über einen längeren Zeitraum und vor allem direkt nach dem Wechsel der Lehrkraft untersuchen. Weiter lassen sich die Ergebnisse nicht auf Schülerpopulationen anderer Schularten übertragen: So steht eine Untersuchung des Zusammenhangs relevanzbezogener Unterrichtsmerkmale mit den Wertüberzeugungen in Mathematik von Schülerinnen und Schülern in Haupt-, Real-, Mittel- oder Berufsschulen noch aus.

Zukünftige Forschungsarbeiten könnten außerdem weitere Unterrichtsmerkmale aus Schüler- und Lehrersicht mit den Wertüberzeugungen von Schülerinnen und Schülern in Verbindung bringen, die im Rahmen dieser Studie nicht abgedeckt wurden. Beispielsweise wäre es interessant, zu untersuchen, unter welchen Bedingungen der Einbezug von Alltagsbeispielen eine Rolle für die Wertüberzeugungen der Schülerinnen und Schüler spielt. So könnte es einen Unterschied machen, ob die Lehrkraft die Beispiele aufführt oder aber die Schülerinnen und Schüler selbst (vgl. Pierce und Stacey 2006). Auch eine genauere Untersuchung der Aufgabentypen und -inhalte, die im Mathematikunterricht eingesetzt werden, könnte weiterführende Erkenntnisse zur Wertentwicklung bringen (vgl. z. B. Neubrand et al. 2013): So könnten Mathematikaufgaben im Hinblick auf ihr Potential, die Schülerinnen und Schüler zur Reflexion über die Bedeutung und Anwendbarkeit der Inhalte anzuregen, klassifiziert und mit Wertüberzeugungen in Verbindung gebracht werden.

Die Erkenntnisse aus der vorliegenden Studie könnten darüber hinaus dazu genutzt werden, um im Rahmen von Interventionsstudien die Wirkung bestimmter Unterrichtsstrategien auf die Wertüberzeugungen von Schülerinnen und Schülern in Mathematik genauer zu ergründen.

## Ausblick

Diese Studie liefert Anregungen zur Identifikation von Unterrichtspraktiken, die potenziell zur Förderung der Wertüberzeugungen der Schülerinnen und Schüler im Fach Mathematik beitragen könnten. Sowohl die Lehrkraft und ihr Mathematikunterricht als auch die Klassenkameradinnen und -kameraden spielen eine Rolle für intrinsischen Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung von Schülerinnen und Schülern. Weitere Studien sind jedoch notwendig, um ein vollständiges Bild von der Rolle dieser Bezugspersonen in der Motivationsentwicklung von Schülerinnen und Schülern in Mathematik zu erhalten.

## Anmerkungen

1 Da weniger als 1 % der Varianz in den Wertüberzeugungen der Schülerinnen und Schüler auf Unterschiede zwischen Schulen zurückzuführen ist, wurde die Schulebene für weitere Analysen außer Acht gelassen.

2 Aufgrund des geringen Zeitraums zwischen dem ersten und zweiten Messzeitpunkt wurde auf die Analyse der Effekte der Unterrichtsmerkmale auf die Wertüberzeugungen zu T2 verzichtet.

3 Eine Replikation des Befundmusters der Ergebnisse im Querschnitt mit den Wertüberzeugungen zu T3 als abhängige Variablen bestätigte die Stabilität der berichteten Befunde im Querschnitt (mit den Wertüberzeugungen zu T1).

## Literatur


Altermatt, E. R., & Pomerantz, E. M. (2003). The development of competence-related and motivational beliefs: An investigation of similarity and influence among friends. *Journal of Educational Psychology, 95*(1), 111–123.

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.

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., & Krauss, S. (2009). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung Nr. 83.* Berlin: Max-Planck-Institut für Bildungsforschung.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A Practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289–300.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects – State, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*(1), 37–68.

Chouinard, R., & Roy, N. (2008). Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *British Journal of Educational Psychology, 78*, 31–50.

Clausen, M. (2002). Unterrichtsqualität: Eine Frage der Perspektive? In D. H. Rost (Hrsg.), *Pädagogische Psychologie und Entwicklungspsychologie (Band 29)*. Münster: Waxmann.

Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*(1), 32–47.

Daniels, Z. (2008). Entwicklung schulischer Interessen im Jugendalter. In D. H. Rost (Hrsg.), *Pädagogische Psychologie und Entwicklungspsychologie (Band 69)*. Münster: Waxmann.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives: Psychological and sociological approaches* (S. 75–146). New York: Freeman.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Maciver, D. (1993). Development during adolescence – The impact of stage-environment fit on young adolescents experiences in schools and in families. *American Psychologist, 48*(2), 90 ¬ 101.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121¬138.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), 478–493.

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence, 20*(2), 507–537.

Freudenthal, H. (1968). Why to teach mathematics so as to be useful. *Educational Studies in Mathematics, 1*(1), 3–8.

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663-677. doi:10.1037/edu0000003

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410¬1412.

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*(2), 509–527.

Juvonen, J., Espinoza, G., & Knifsend, C. (2012). The role of peer relationships in student academic and extracurricular engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Hrsg.), *Handbook of research on student engagement* (S. 387–492). New York: Springer.

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Development, 78*(4), 1186–1203.

Krapp, A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie in Erziehung und Unterricht, 45*(3), 185–201.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education, 33*(1), 27–50.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231–251.

Kunter, M., & Voss, T. (2013). The model of instructional quality in COACTIV: A Multicriteria analysis. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers* (S. 97–124). New York: Springer.

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*(5), 494–509.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau, 58*(2), 103–117.

Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Analyse von Lernumwelten. *Zeitschrift für Pädagogische Psychologie, 20*(1), 85–96.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*(3), 203–229.

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*(2), 106–124.

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B. O., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764–802.

Michelsen, C., & Sriraman, B. (2009). Does interdisciplinary instruction raise students' interest in mathematics and the subjects of the natural sciences? *ZDM – The International Journal on Mathematics Education, 41*(1-2), 231–244.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide: Seventh edition*. Los Angeles, CA: Muthén & Muthén.

Nelson, R. M., & DeBacker, T. K. (2008). Achievement motivation in adolescents: The role of peer climate and best friends. *Journal of Experimental Education, 76*(2), 170–189.

Neubrand, M., Jordan, A., Krauss, S., Blum, W., & Löwen, K. (2013). Task analysis in COACTIV: Examining the potential for cognitive activation in German mathematics classrooms. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers* (S. 125–144). New York: Springer.

Osborne, J., & Dillon, J. (2008). *Science Education in Europe: Critical Reflections*. London: The Nuffield Foundation.

Parker, J. G., Rubin, K. H., Erath, S. A., Wojslawowicz, J. C., & Buskirk, A. A. (2006). Peer relationships, child development, and adjustment: A developmental psychopathology perspective. In D. Cicchetti & D. J. Cohen (Hrsg.), *Developmental Psychopathology: Theory and Method (Vol. 1)*. Hoboken, NJ: John Wiley & Sons, Inc.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Hrsg.), *Handbook of research on student engagement* (S. 365–386). New York: Springer.

Pierce, R. U., & Stacey, K. C. (2006). Enhancing the image of mathematics by association with simple pleasures from real world contexts. *ZDM – The International Journal on Mathematics Education, 38*(3), 214–225.

Rakoczy, K., Klieme, E., & Pauli, C. (2008). Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation. *Zeitschrift für Pädagogische Psychologie, 22*(1), 25–35.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Vol. 1)*. Thousand Oaks: SAGE Publications, Incorporated.

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology, 10*(2), 321–352.

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *Elementary School Journal, 100*(5), 443–471.

Ryan, A. M. (2000). Peer groups as a context for the socialization of adolescents' motivation, engagement, and achievement in school. *Educational Psychologist, 35*(2), 101–111.

Ryan, A. M. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development, 72*(4), 1135–1150.

Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel & A. Wigfield (Hrsg.), *Handbook of Motivation at School* (S. 197–222). New York: Routledge.

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in 2-level models. *Sociological Methods & Research, 22*(3), 342–363.

Steinmayr, R., & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica, 56*(4), 195–211.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777.

Trautwein, U., Nagengast, B., Marsh, H. W., Gaspard, H., Dicke, A.-L., Lüdtke, O., & Jonkmann, K. (2013). Expectancy-value theory revisited: From expectancy-value theory to expectancy-valueS theory? In D. M. McInerney, H. W. Marsh, & R. G. Craven (Hrsg.), *Theory driving research: New wave perspectives on self-processed and human development* (S. 233–249): Information Age Publishing, Incorporated.

Turner, J. C., & Meyer, D. K. (2009). Understanding motivation in mathematics: What is happening in classrooms? In K. Wetzel & A. Wigfield (Hrsg.), *Handbook of motivation at school* (S. 527–552). New York: Routledge.

Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., Matos, L., & Lacante, M. (2004). Less is sometimes more: Goal content matters. *Journal of Educational Psychology, 96*(4), 755–764.

Vorhölter, K. (2009). *Sinn im Mathematikunterricht: Zur Rolle von Modellierungsaufgaben bei der Sinnkonstruktion von Schülerinnen und Schülern. Studien zur Bildungsgangforschung (Band 27)*. Leverkusen: Barbara Budrich.

Vorhölter, K., & Vollstedt, M. (2012). Zur theoretischen Konzeption und zu den Möglichkeiten der unterrichtspraktischen Umsetzung der Sinnkonstruktion. In W. Blum, R. Borromeo Ferri, & K. Maaß (Hrsg.), *Mathematikunterricht im Kontext von Realität, Kultur und Lehrerprofessionalität* (S. 148–156). Wiesbaden: Vieweg+Teubner Verlag.

Wang, M. T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology, 48*(6), 1643–1657.

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Development, 75*(5), 1556–1574.

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1–35.

Wigfield, A., & Eccles, J. S. (2002). Students' motivation during the middle school years. In J. Aronson (Hrsg.), *Improving academic achievement: Impact of psychological factors on education* (S. 159–184). Amsterdam: Academic Press.

Wigfield, A., Tonks, S., & Klauda, S. (2009). Expectancy-value theory. In K. Wetzel & A. Wigfield (Hrsg.), *Handbook of motivation at school* (S. 55–75). New York: Routledge.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg (Hrsg.), *Handbook of child psychology (6th edition, vol. 3)* (S. 933–1002). Hoboken, New Jersey: John Wiley & Sons, Inc.

Willems, A. S. (2011). Bedingungen des situationalen Interesses im Mathematikunterricht: Eine mehrebenenanalytische Perspektive. *Empirische Erziehungswissenschaft (Band 30)*. Münster: Waxmann.

# 5

# Study 2

**Short intervention, sustained effects:**

**Promoting students' mathematics-related competence beliefs, effort, and achievement**

Brisson, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Nagengast, B., & Trautwein, U. (2017). Short intervention, sustained effects: Promoting students' math competence beliefs, effort, and achievement. *American Educational Research Journal, 54*(6), 1048-1078. DOI: 10.3102/000283121771608.

**Abstract**

The present study investigated the effectiveness of two short relevance interventions (writing a text or evaluating quotations about the utility of mathematics) using a sample of 1916 students in 82 mathematics classrooms in a cluster randomized controlled experiment. Short-term and sustained effects (six weeks and five months after the intervention) of the two intervention conditions on students' competence beliefs (self-concept, homework-related self-efficacy), teacher-rated individual effort, and standardized test scores in mathematics were assessed. Hierarchical linear regression analyses showed that students' homework-related self-efficacy was higher in both intervention groups six weeks and five months after the intervention compared to the control condition. Students' self-concept, teacher-rated effort, and achievement in mathematics were promoted through the quotations condition, partly in the long term.

**Keywords:** expectancy-value theory · relevance intervention · competence beliefs · effort · mathematics achievement

## Introduction

How can secondary school students be supported to become more self-confident, hard-working, and successful in mathematics? To foster student motivation and performance, especially in science, technology, engineering, and mathematics (STEM) subjects, researchers and educational stakeholders promote relevance-enhanced teaching (e.g., Davis & McPartland, 2012; Osborne & Dillon, 2008). Indeed, yearlong teaching programs systematically emphasizing connections between mathematical learning material and career opportunities have been found to raise students' mathematics grades (Woolley, Rose, Orthner, Akos, & Jones-Sanpei, 2013), and shorter interventions using writing assignments about the personal relevance of STEM subjects have been shown to improve perceived utility of and interest in STEM (e.g., Hulleman & Harackiewicz, 2009).

These findings are promising, but only little is known about the potential of short relevance interventions implemented in school classrooms (for an exception, see Hulleman & Harackiewicz, 2009). First, there is a need for comparative studies to investigate the relative strength of different intervention approaches. To this end, successful intervention strategies could be combined or added with new features to create various treatment conditions. Second, the majority of studies on the effects of classroom-based relevance interventions focused mainly on the focal construct (value beliefs) and achievement as outcomes. The impact of relevance interventions on students' competence beliefs and effort, however, has not yet been investigated in school classroom settings. Concerning treatment effects on performance, students' grades or exam scores, but no standardized test scores, have been used as achievement measures producing inconsistent findings (e.g., Hulleman, Godes, Hendricks, & Harackiewicz, 2010, study 2; Woolley et al., 2013). Besides, no other outcomes than grades have so far represented the teachers' perspective in the evaluation of relevance interventions.

To shed light on these research gaps, we used data from the Motivation in Mathematics (MoMa) study in which two different relevance interventions (one adapted from previously used approaches and one novel one) were implemented in 82 mathematics classrooms in Grade 9 using a cluster randomized controlled study design. Prior analyses have found these interventions to improve students' value beliefs of mathematics (Gaspard, Dicke, Flunger, Brisson, et al., 2015) and students' self-reported effort (Gaspard et al., 2016). The present study analyzed and compared the short-term and sustained effects of the same treatments on further outcomes neglected in previous classroom-based relevance experiments: on students' self-concept, homework-related self-efficacy, teacher-rated effort, and standardized test scores in mathematics.

### The importance of perceived utility value in mathematics

The Eccles et al. (1983) expectancy-value theory (EVT) is a powerful framework highlighting the importance of students' perceived utility value in determining students' achievement-related behaviors and performance (Eccles & Wigfield, 2002). According to EVT, students perceive high levels of utility value when they believe that engaging in an academic task will help them reach their personal goals. With regards to intrinsic and extrinsic motivation—two motivational concepts referring to either doing an activity for inherent satisfaction or to reach some separable outcome (self-determination theory, e.g., Ryan & Deci, 2000)—the utility value component defined in EVT simultaneously comprises both intrinsic and extrinsic reasons for putting effort in a task (Eccles, 2005). More precisely, task completion is valued because the outcome of the task is expected to serve another end; this goal, however, may be personally meaningful to the student. Supporting students to relate the learning contents to their personal goals and to thus link intrinsic and extrinsic reasons for task engagement seems a promising approach for classroom motivational interventions (e.g., Trautwein et al., 2013).

Numerous empirical studies underline that it is beneficial for students' motivation, behavior, and performance when students perceive the learning contents to be useful (for overviews, see Roeser, Eccles, & Sameroff, 2000; Wigfield & Cambria, 2010; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). In mathematics, students reporting high levels of utility value also show high levels of competence beliefs (for instance, self-efficacy and ability perceptions), effort, and achievement (e.g., Cole, Bergin, & Whittaker, 2008; Eccles & Wigfield, 1995; Husman & Hilpert, 2007). However, studies on developmental trajectories demonstrate that students' mathematics-related utility value beliefs are decreasing continuously through secondary school (e.g., Chouinard, Karsenti, & Roy, 2007; Chouinard & Roy, 2008; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). In line with these findings, interviews have shown that secondary school students have a hard time coming up with concrete examples for the utility of mathematical knowledge in real-life situations (Harackiewicz, Hulleman, Rozek, Katz-Wise, & Hyde, 2010).

Researchers have therefore examined how students' perceived utility value can be promoted and found relevance-enhanced teaching approaches to bear a huge potential in fostering STEM-related student outcomes in both laboratory and natural learning settings (for overviews, see Durik, Hulleman, & Harackiewicz, 2015; Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016; Yeager & Walton, 2011). Two types of strategies have been employed to convey the relevance of STEM subjects to students: (a) providing information about the utility of the learning material, for instance, for daily life (e.g., Durik & Harackiewicz, 2007, study 2); (b) having students generate arguments for the utility of the learning material themselves (e.g., Hulleman & Harackiewicz, 2009). Results concerning the effectiveness of these intervention strategies, however, vary across different types of settings (laboratory vs. classroom), outcomes, and students' prerequisites (see Durik, Hulleman, et al., 2015). In a series of lab studies (Canning & Harackiewicz, 2015, studies 1 and 2; Durik & Harackiewicz, 2007, study 2; Durik, Shechter, Noh, Rozek, & Harackiewicz, 2015, study 1; Hulleman et al., 2010, study 1; Shechter, Durik, Miyamoto, & Harackiewicz, 2011, study 1), both strategies have been shown to raise undergraduates' perceived utility of and interest in a multiplication technique. Furthermore, the provision of utility information promoted students' involvement, effort, competence valuation, perceived competence, and test scores when applying the same technique—in particular for high-achievers. For low-achievers, a combination of both strategies has been found to increase students' perceived utility of and interest in the multiplication technique as well as test scores when applying the technique (Canning & Harackiewicz, 2015, study 2).

Fewer studies intervened on students' utility value of STEM subjects in real-life classroom settings, but their success is compelling: Providing information about the utility of mathematical learning contents for career opportunities has been found to foster secondary school students' math grades (Woolley et al., 2013). Having students generate arguments for the relevance of specific topics in science or psychology courses promoted students' utility value, interest, success expectancies, and—partially—students' grades or exam scores, especially for students with low actual or perceived competence (Hulleman et al., 2010, study 2; Hulleman & Harackiewicz, 2009; Hulleman, Kosovich, Barron, & Daniel, 2017, study 2). An overview of the central characteristics of these classroom-based studies (setting, sample size, intervention, evaluation design, and results) is provided in Table 1. Drawing from these studies, we created two relevance interventions including new features with regards to the focus, the strategies, and the level of the interventions, and compared their short-term and sustained effects on previously neglected outcomes, including different perspectives (students and teachers).

**Characteristics of the MoMa interventions**

In previous school interventions in STEM subjects, students typically looked into the relevance of specific learning topics for their lives using numerous writing assignments or teacher-led lessons (see Table 1). However, instead of concentrating on topic-specific relevance, students in the MoMa interventions had to reflect on the personal relevance of mathematics as a broader domain, in particular for future education and career pathways. This approach aims to support students' math investment over and above the topic currently dealt with in class—even when given a rather short intervention (cf., correlational and experimental research on the importance of students' school and professional goals for their math investment, e.g., Peetsma & van der Veen, 2011; Schuitema, Peetsma, & van der Veen, 2014).

In addition, the MoMa interventions integrated previous successful intervention approaches, namely presenting and self-generating utility arguments (e.g., Durik & Harackiewicz, 2007, study 2; Hulleman & Harackiewicz, 2009), into one approach. Combining different interventions may have additive effects if the interventions depend on different mechanisms (Yeager & Walton, 2011). Self-generating utility arguments in individual writing assignments enables students to make personalized connections with the learning material (Hulleman et al., 2017). The personalization of the intervention message, in turn, has been found to be crucial for the meaningfulness and effectiveness of educational interventions (Walton, 2014; Yeager & Walton, 2011). However, as students might lack concrete examples of the utility of mathematics in real-life situations (Harackiewicz et al., 2010) generating utility arguments in individual essays without any preparation (e.g., Hulleman & Harackiewicz, 2009) might be a difficult task for them to do. Presenting some examples for the usefulness of mathematics, for instance, for specific education and career pathways, might help students in reflecting about their own personal relevance of mathematics in a more productive way. In addition, discussing

Table 1
*Overview of relevance interventions in STEM subjects in the classroom*

| Study | Setting | Sample size | Intervention | | | Evaluation design | Reported results |
|---|---|---|---|---|---|---|---|
| | | | Level | Instructor | Design | | |
| Hulleman & Harackiewicz (2009) | High school Grade 9<br>Subject: science (biology, integrated sc., physics) | *n* = 262<br>E: *n* = 136<br>C: *n* = 126 | Student | Research assistant, teacher | 8 essays in 1 semester:<br>E: describe utility of course topic to one's life<br>C: summarize course topic | SQ before first and after last essay<br>end-of-semester grade (1 week after last essay)<br>no follow-up | No main effects<br>Effects on interest in science and grades moderated through success expectations<br>No effects on interest in science-related courses and careers<br>No moderation through gender and race |
| Hulleman, Godes, Hendricks, and Harackiewicz (2010), Study 2 | College Introductory psychology course | *n* = 318<br>E1: *n* = 78<br>E2: *n* = 82<br>C1: *n* = 78<br>C2: *n* = 80 | Student | Research assistant | 2 essays in 3 weeks:<br>E1: describe relevance of course topic in a letter to a significant person in one's life<br>E2: discuss relevance of media report for course topic<br>C1: summarize course topic<br>C2: discuss how abstract of scientific article expands upon course topic | SQ before 1st and after 2nd essay<br>end-of-year grade (3 weeks after last essay)<br>no follow-up | Main effects of E1 and E2 on situational interest<br>Effects of E1 and E2 on utility value and maintained interest moderated through initial performance<br>No effects on grades |
| Hulleman, Kosovich, Barron, and Daniel (2017), Study 2 | University Introductory psychology course | *n* = 357<br>E1: *n* = 116<br>E2: *n* = 122<br>C: *n* = 119 | Student | Teacher (online) | 2 essays (after 1st and 2nd exam):<br>E1, E2: relate course material to one's life<br>E2 (enhanced intervention): (after 1st exam) create implementation intentions for relating course material to one's life, (after 2nd exam) reflect on self-regulation strategies<br>C: summarize course topic | SQ before 1st and about six weeks after 2nd essay<br>exam scores throughout the semester<br>final course grade | Main effects of E1 and E2 on success expectancy and grade<br>Effects of E1 and E2 on interest and success expectancy moderated through initial performance<br>Effects of E1 and E2 on final exam scores moderated through success expectancy and initial performance; 3-way-interactions with initial performance, gender<br>No effects on utility value or cost |
| Woolley, Rose, Orthner, Akos, and Jones-Sanpei (2013) | Middle school Grades 6-8<br>Subject: mathematics | *n* = ~6,500<br>E: *n* = 3,295<br>C: *n* = ~3,200 | School | Teacher | E: follow 10 teacher-led math lessons per year in Grades 6-8 treating career relevant examples and math problems (standard curriculum)<br>C: regular instruction | end-of-year grades in Grades 3-5 and in Grades 6-8<br>no follow-up | Main effects on grades in Grades 7 & 8<br>No effects on grades in Grade 6 |

*Note*. sc = science; n = number; E = experimental group; C = control group; SQ = student questionnaire.

occupations in which general math knowledge and analytic skills are needed might create a moment of sudden insight for students—in particular, when the need for mathematics is not very obvious (e.g., for studying social sciences). This might help change the way students think about the relevance of learning mathematics at school (cf., Walton, 2014). We expected that the effectiveness of the first MoMa intervention condition, i.e., writing a text about the personal relevance of mathematics (Hulleman & Harackiewicz, 2009), would benefit from a preceding input on the usefulness of mathematical skills.

Another way to possibly enhance the effectiveness of social-psychological interventions is the use of contextually appropriate anecdotes or quotations from peers such as older students about situations in which they needed mathematical knowledge (Yeager & Walton, 2011). This assumption is supported by a social cognition perspective as found in social learning theory (Bandura, 1977), possible-selves theory (Markus & Nurius, 1986), and identity-based motivation (Oyserman & Destin, 2010), which postulate that students can learn from persons they identify with. Accordingly, slightly older students or young adults describing the usefulness of mathematics in their lives could help students imagine a potential future identity and the importance of mathematical skills in developing this identity. As statements from interviews provide personal and authentic utility information, they might be an effective tool to encourage students' personal reflection about the relevance of mathematics (see Harackiewicz, Rozek, Hulleman, & Hyde, 2012 who used a similar approach as part of a more comprehensive motivation intervention in STEM subjects). Having students evaluate quotations from slightly older peers about the relevance of mathematics in the second MoMa intervention condition was thus aimed at supporting students' own valuing of mathematics.

Compatibility with students' natural learning environment is an important precondition for the effectiveness of classroom-based interventions. Previous relevance interventions in STEM subjects were mainly conducted at the student level (see Table 1). As students are typically taught together in classes, however, intervening at the classroom level would come closer to the natural learning setting. At the same time, class-level interventions allow for students' active participation, for instance, in discussions about the relevance of mathematics. This might help in triggering personal reflection and thus increase treatment effects. As an additional advantage, between-class experimental designs allow for a more precise estimation of the intervention effects: They bear a reduced risk of diffusion effects which occur in within-class experimental designs when classmates randomized into different intervention conditions interact with each other (Craven, Marsh, Debus, & Jayasinghe, 2001).

Lastly, we also evaluated the effectiveness of the interventions more broadly than the studies presented in Table 1. More precisely, research is missing investigating direct treatment effects of classroom-based relevance interventions on motivational, behavioral, and achievement outcomes simultaneously. Findings so far considered students' grades (all studies shown in Table 1), interest (studies by Hulleman et al.), utility value (Hulleman et al., 2010; 2015, study 2), and cost and success expectancy (Hulleman et al., 2015, study 2). However, further motivational outcomes such as students' self-concept and self-efficacy, behavioral outcomes such as effort, and standardized performance measures have been neglected in previous research. As all previous outcome measures with the exception of grades were measured using students' self-reports, the teacher's perspective has not yet been considered in the evaluation of the effectiveness of relevance interventions. Moreover, the sustainability of the intervention effects through the use of a follow-up has so far only been investigated for performance (Hulleman et al., 2017, study 2; Woolley et al., 2013).

**Competence beliefs, effort, and test scores: understudied outcomes of classroom-based relevance interventions**

A closer examination of the Eccles' et al. (1983) expectancy-value theory suggests a range of educational outcomes that could be affected by relevance interventions. First of all, EVT assumes students' value beliefs to be positively interrelated (Eccles & Wigfield, 2002), which implies that promoting students' utility value may also foster other value beliefs (see Gaspard, Dicke, Flunger, Brisson, et al., 2015, for the effects of the MoMa interventions on value beliefs) . Furthermore, according to EVT, students' utility value is closely associated with students' competence beliefs and predicts achievement-related behaviors (e.g., effort) and test performance (Eccles & Wigfield, 2002; Wigfield et al., 2006)—outcomes which are understudied when analyzing the effectiveness of relevance interventions in secondary schools.

If students are aware of the utility of a subject for attaining their personal goals, they may be ready to tackle related tasks intensely and thereby discover their academic potential in a domain (see Hulleman et al., 2017). They may also be willing to put in more effort, thus, positively engaging in learning (e.g., Reschly & Christenson, 2012). Hence, pondering over the relevance of the learning material could promote students' academic self-concept, self-efficacy, and effort. Students' academic self-concept is a domain-specific competence belief referring to how students evaluate their abilities in an academic domain (Eccles & Wigfield, 2002). Students' math self-concept has been found to be a strong predictor of students' interest, effort, persistence, choice of task difficulty, course choice, and performance in mathematics (e.g., Denissen, Zarrett, & Eccles, 2007; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009). Students' self-efficacy is a task-specific competence belief assessing students' confidence in their ability to successfully accomplish a specific task like their mathematics homework (Bandura, 1994). Students' homework-related self-efficacy has been shown to influence students' homework-related value beliefs as well as homework effort and compliance in mathematics (Trautwein, Lüdtke, Schnyder, & Niggli, 2006)—behaviors which, in turn, impact mathematics performance (e.g., Zimmerman & Kitsantas, 2005).

Supporting the assumptions made in EVT, numerous non-experimental studies have shown positive associations of secondary school students' utility value beliefs with their self-concept or homework-related self-efficacy in mathematics (e.g., Chouinard et al., 2007; Eccles & Wigfield, 1995; Husman & Hilpert, 2007; Jacobs et al., 2002; Trautwein & Lüdtke, 2009) as well as with effort in mathematics (e.g., Chouinard et al., 2007; Cole et al., 2008; Trautwein, Lüdtke, Kastens, & Köller, 2006; Trautwein, Lüdtke, Schnyder, et al., 2006). In addition, in laboratory experiments, subgroups of students (e.g., low-achievers) were more confident in applying a new multiplication technique correctly and put more effort in using the technique after reading about its utility (Durik & Harackiewicz, 2007, study 2; Shechter et al., 2011, study 1). Similarly, a classroom intervention during which undergraduate students collected arguments about the personal relevance of various topics in introductory psychology fostered low-achievers' expectancies to succeed in the class (Hulleman et al., 2017, study 2). However, the effects of relevance interventions conducted in secondary school classrooms on students' domain-specific self-concept, task-specific self-efficacy, and effort as well as the sustainability of such effects have not yet been investigated.

Furthermore, relevance interventions could promote students' test performance. Yet, whereas laboratory-based relevance experiments have been found to foster students' test scores (e.g., Canning & Harackiewicz, 2015, studies 1 and 2; Durik & Harackiewicz, 2007, study 2), classroom-based intervention studies have only investigated students' grades or exam results as achievement outcomes so far; these analyses yielded inconsistent results, namely either main effects (Woolley et al., 2013), or moderated effects (e.g., Hulleman & Harackiewicz, 2009; Hulleman et al., 2017, study 2), or no effects (Hulleman et al., 2010, study 2) on grades or exam scores. These mixed results might in part be due to teachers' subjective grading practices (e.g., McMillan, 2001). Consequently, there is a need to analyze whether classroom-based relevance interventions promote achievement measured by standardized test scores.

**The present study**

In the present study, we investigated the short-term and sustained effects of two short relevance intervention conditions (quotations, text) implemented at the classroom level on ninth-grade students' competence beliefs, teacher-rated effort, and test scores in mathematics compared to a control group. Based on previously established approaches, students were first presented arguments for the usefulness of mathematics and then reflected on the personal usefulness of mathematics in an individual writing assignment. Drawing on a social cognition perspective (Bandura, 1977; Markus & Nurius, 1986; Oyserman & Destin, 2010) and prior intervention approaches (Harackiewicz et al., 2012), students in the quotations condition commented on interview statements from young adults about the relevance of mathematics. Adapted from a successful strategy first tested by Hulleman et al. (2009; 2010), students in the text condition wrote texts about the personal relevance of mathematics. We included a broad range of important outcomes, namely students' self-concept, homework-related self-efficacy, effort, and standardized test scores in mathematics. As students' effort is observable (e.g., Fredricks, Blumenfeld, & Paris, 2004), teachers rated individual students' effort in the current study, thereby including teachers' perspective on the effectiveness of the interventions and going beyond previous investigations concerning student-reported effort (Gaspard et al., 2016).

To learn about the sustainability of the intervention effects, we used a follow-up design evaluating treatment effects six weeks and five months after the interventions.

Prior analyses with the same data set showed that students' utility value was fostered through both intervention conditions for at least five months, and that students' other value beliefs of mathematics (attainment and intrinsic value) except for cost were promoted to different degrees (Gaspard, Dicke, Flunger, Brisson, et al., 2015). Furthermore, the quotations condition had stronger effects on students' self-reported effort than the text condition (Gaspard et al., 2016). Grounded on EVT (Eccles & Wigfield, 2002), findings from correlational studies (e.g., Chouinard et al., 2007; Cole et al., 2008; Trautwein, Lüdtke, Kastens, et al., 2006) and experimental research (e.g., Durik & Harackiewicz, 2007, study 2; Hulleman et al., 2017, study 2), we hypothesized students' self-concept, homework-related self-efficacy, teacher-rated effort, and test scores in mathematics to be promoted through both intervention conditions. Due to lack of empirical evidence, no hypotheses were formulated concerning the stability of the treatment effects and the comparative strength of the two intervention conditions.

## Method

### Sample and data collection

Data were gathered in the project "Motivation in Mathematics" (MoMa) in 82 ninth grade mathematics classrooms from 25 academic track schools ("Gymnasium") in the German state of Baden-Württemberg. The sample size was based on a power analysis for a multi-site cluster randomized trial indicating a power of $\beta$ = .73 to detect an effect of $\delta$ = 0.20 per intervention condition compared to the control condition (see Gaspard, Dicke, Flunger, Brisson, et al., 2015, for more information). In the present sample, mathematics was taught as one comprehensive subject including different domains such as algebra, geometry, or calculus during four compulsory lessons per week. There was no further tracking of students in mathematics courses within school. Homework assignments in mathematics were common in all but one class (98.8 %). A total of 1978 students with active parental consent participated in the study, corresponding to a participation rate of 96.0 %. Sixty-two students absent during the day of the intervention were excluded from the analyses, yielding a sample of 1916 students (53.3 % female; mean age at the start of the study: $M$ = 14.41 years, $SD$ = 0.57; mean SES/ISEI[1]: $M$ = 65.24, $SD$ = 16.21). The large majority of students were Caucasian, and students with an immigrant background (21.2 % with at least one parent born outside Germany) came from predominantly Western countries and were Caucasian.

Data collections took place from September 2012 to March 2013 and were administered by trained researchers. Students in the intervention conditions completed questionnaires before the intervention (pretest = T1) as well as six weeks (posttest = T2) and five months (follow-up = T3) after the intervention. Students in the waiting control group completed the same questionnaires at the same time points but did not receive any intervention before T3. Students' competence beliefs and effort were measured at all three time points. Students' mathematics achievement was measured in the beginning of the school year and at the follow-up. Students' mathematics-related utility value was also measured at all three time points and will be reported to give an account of how it was associated with the outcome variables and affected by the interventions (see also Gaspard, Dicke, Flunger, Brisson, et al., 2015). All 82 classes fully completed all waves of data collections.

**The MoMa relevance interventions**

In the beginning of the study, all 73 participating teachers and their classes[2] were randomly assigned within their schools to one of the three study conditions (quotations: 25 classes, 561 students; text: 30 classes, 720 students; waiting control group: 27 classes, 635 students[3]). Before the first data collection, teachers participated in an information session about the design and the theoretical background of the study. To gain teachers' trust in the project and to avoid spillover effects (Craven et al., 2001), teachers in the waiting control group were informed that their classes would also receive the intervention after the last data collection, and that they were not supposed to ask their colleagues in the experimental groups about the contents of the intervention. Teachers in the experimental groups were not informed whether their classes had been assigned to the quotations condition or to the text condition.

After students in all treatment conditions had completed the pretest, students in the intervention conditions received a 90-minute standardized relevance intervention led by five trained researchers in class and followed by two short intervention boosters to be completed at home after two and three weeks. To control for implementation fidelity, researchers recorded the actual procedure of each intervention in the minutes. Every researcher conducted 8 to 13 interventions with roughly equal distribution between the two intervention conditions.

The interventions were designed combining previously tested strategies, namely the presentation and the self-generation of relevance arguments (e.g., Canning & Harackiewicz, 2015, study 2), with newly developed features. As a result, the interventions consisted of a psychoeducational presentation and an individual writing assignment differing by condition. High initial competence beliefs have been shown to be a prerequisite for appreciating relevance information (Canning & Harackiewicz, 2015; Durik, Hulleman, et al., 2015; Durik, Shechter, et al., 2015, study 2). As a confidence booster, students were informed about research results concerning the importance of effort, different interpretations of achievement-related experiences, and frame of reference effects in school classrooms (see Marsh, 2005; Wigfield et al., 2006) in the first part of the presentation. The second and main part of the presentation dealt with the usefulness of mathematics as a broader domain for future education, career opportunities, and leisure time activities.

After the presentation, students completed individual writing assignments differing by condition. Based on theories of social cognition that assume that students can learn from persons they identify with (e.g., Bandura, 1977; Markus & Nurius, 1986; Oyserman & Destin, 2010), students in the quotations condition were encouraged to reflect on the personal relevance of mathematics by reading six interview quotations from older students or young adults who describe the usefulness of mathematics to their lives. Covering a broad range of real-life situations, the quotations stem from a preceding interview study in which thirty persons (ranging from high school graduates to working adults) were asked to describe personal situations where they needed mathematical skills. During the intervention, the students were asked to evaluate the relevance of these quotations to their own lives by responding to a set of questions (see Appendix, Part A, for sample quotations and the work assignments). Students in the text condition were asked to collect arguments for the personal relevance of mathematics to their current and future lives, and to then write a coherent text detailing their notes. This task was adapted from relevance interventions by Hulleman et al. (2009; 2010) by switching the focus of the assignment from specific course topics treated in class to mathematics as a domain (see Appendix, Part A, for the instruction).

At the end of the intervention, students received a portfolio including two short intervention boosters to be filled out at home one and three weeks after the intervention session, respectively. In the first booster, students were asked to summarize what they remembered from their individual writing assignments in class. The second booster differed by condition and corresponded to the type of individual assignment dealt with in class (quotations: reflection on given relevance information; text: self-generation of relevance arguments). Students in the quotations condition were asked to choose one out of several arguments about the relevance of mathematics provided on a website (www.dukannstmathe.de) and to describe why it was convincing to them. Students in the text condition were asked to explain why mathematics was useful to a person they knew.

Students in classes in the waiting control condition did not follow any presentation or do any individual writing assignments. However, they received the more successful intervention approach after the last measurement point.

### Measures

**Mathematics-related competence beliefs.** Students' mathematics-related competence beliefs were assessed with a student questionnaire using four-point Likert type scales ranging from 1 (*completely disagree*) to 4 (*completely agree*) that were adapted from previous studies (e.g., Baumert, Gruehn, Heyn, Köller, & Schnabel, 1997; Prenzel et al., 2006). Mathematics-related self-concept was measured with five items (e.g., "I am good at math.", $\alpha$ = .93). The homework-related self-efficacy scale consisted of four items (e.g., "When I try hard, I can solve my math homework correctly.", $\alpha$ = .76).

**Effort in mathematics.** Teachers rated individual students' mathematics-related effort by responding to the item "This student works thoroughly on all of his/her math tasks and homework assignments." on a four-point Likert type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*).

**Achievement in mathematics.** Students' results from a curriculum-based standardized test assessing mathematical knowledge in the state of Baden-Württemberg in the beginning of Grade 9 served as an initial measure of mathematics performance. The test assessed students' competencies in the mathematical domains of algebra, geometry, and probability calculus with 38 math problems. The math problems focused on three aspects of math proficiency: numbers and algorithms, space and shapes, linking and modeling (38 questions; assessed by percent correct). At the follow-up, students completed a 3-minute normed speed test, which measured students' fluency of solving typical mathematical operations (50 questions; max. number of points: 50) (Schmidt, Ennemoser, & Krajewski, 2013). Validity studies showed that this short speed test is a very good proxy for students' achievement in longer assessments using standardized, curriculum-based test in mathematics (Ennemoser, Krajewski, & Schmidt, 2011; Schmidt et al., 2013). The internal consistency of the test was good (Cronbach's $\alpha = .89$).

**Mathematics-related utility value.** Students' mathematics-related utility value beliefs was measured through student ratings using a four-point Likert type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). A comprehensive utility value scale consisting of twelve items (e.g., "I will often need math in my life.") out of a newly developed value instrument was used (Gaspard, Dicke, Flunger, Schreier, et al., 2015). The scale showed good internal consistency (Cronbach's $\alpha = .84$) (for more details, see Gaspard, Dicke, Flunger, Schreier, et al., 2015).

### Statistical analyses

**Multilevel regression analyses.** In order to test the treatment effects on students' competence beliefs, teacher-rated effort, and achievement, two-level linear regression analyses[4] were computed with Mplus (Version 7; Muthén & Muthén, 1998-2012) for each of the outcome variables. Separate multilevel regression analyses were carried out using students' competence beliefs and teacher-rated effort at T2 and T3 as well as test scores at T3 as outcomes and two dummy variables indicating the treatment (quotations, text) as class-level predictors. Each outcome variable was regressed on the intervention conditions at the class level, the control condition being the reference group. In line with the recommended procedure to test intervention effects in cluster randomized trials (Raudenbush, 1997), initial values of the respective outcome variables were used as covariates both at the student level and at the class level. To account for contextual effects, all effects on the respective outcomes were freely estimated at both levels (Korendijk, Hox, Moerbeek, & Maas, 2011; Marsh et al., 2009). Covariates were added to the models using group-mean centering at the student level (Enders & Tofighi, 2007) and manifest aggregation at the class level (Marsh et al., 2009).

**Effect sizes.** Before running the analyses, all continuous (but not the dichotomous) variables were standardized. Consequently, the regression coefficients of the dummy variables can be directly interpreted as measures of the class-level effect sizes of the intervention conditions on the outcomes as compared to the control condition (Marsh et al., 2009; Tymms, 2004).

**One-tailed vs. two-tailed tests.** To evaluate the statistical significance of the treatment effects, the use of two-tailed tests is recommended particularly if the literature does not support any directional hypotheses (e.g., Howell, 2012). Yet, given our directional a priori hypotheses, the significance of the treatment effects was tested on the basis of one-tailed tests with an $\alpha$-

level of 5 %. This testing procedure additionally improves the power to detect small treatment effects at the class level (Stevens, 2012).

**Missing data.** Missing data ranged from 2.3 % to 19.6 % for the outcome variables (see Table 2). Based on suggestions for the treatment of missing values by Graham (2009), the full information maximum likelihood method integrated in Mplus was used to deal with missing data. To make the assumption of missing at random more plausible, correlations of three auxiliary variables (students' gender, pretest cognitive ability score assessed with a figural cognitive ability test by Heller & Perleth, 2000, and end-of-year math grade in Grade 8) with the predictor variables were included in the models at both levels (Enders, 2010). The auxiliaries' and predictors' residuals were also included in the models at both levels.

**Implementation fidelity.** To account for implementation fidelity, analyses were run with two types of samples: a) including all classes participating in the interventions, and b) excluding two classes in which deviations from the intervention manual had been recorded in the minutes. Deviations occurred in two classes in the text condition: In one class, the initial presentation had to be held without any projector due to technical problems; in the other class, the researcher conducting the intervention noted that students were reluctant to participate in the intervention and, in particular, did not work quietly on their individual writing assignments. A comparison of the results showed no noteworthy differences, which is why all classes were included in the final analyses.

## Results

### Descriptive statistics, randomization check, and effects on mathematics-related utility value

Before analyzing treatment effects, the descriptive statistics (see Table 2) and the inter-correlations of all outcome variables including utility value beliefs (see Table 3) were calculated at all measurement points. As a randomization check, the differences in the pretest means for students' perceived utility value, competence beliefs, teacher-rated effort, and achievement between the three study conditions were tested for statistical significance. No statistically significant differences between the conditions emerged, based on two-tailed Wald-$\chi^2$-tests (Bakk & Vermunt, 2016) with an α-level of 5 % (utility value: $\chi^2$ (2) = 0.79, $p$ = .675; self-concept: $\chi^2$ (2) = 0.88; $p$ = .643; homework-related self-efficacy: $\chi^2$ (2) = 3.73; $p$ = .155; teacher-rated effort: $\chi^2$ (2) = 5.01; $p$ = .082; test score: $\chi^2$ (2) = 1.51; $p$ = .470). Concerning intervention effects on the focal construct, students' mathematics-related utility value beliefs, analyses revealed a significant promotion through both the quotations condition (posttest: $\beta$ = .30; $p$ < .000; follow-up: $\beta$ = .26; $p$ < .001) and the text condition (posttest: $\beta$ = .14; $p$ = .011; follow-up: $\beta$ = .16; $p$ = .004) (see Gaspard, Dicke, Flunger, Brisson, et al., 2015).

Table 2

*Descriptive statistics of the sample characteristics and the study variables per intervention condition*

| | | Quotations (561 students, 52.8 % female) | | | | Text (720 students, 52.4 % female) | | | | Control group (635 students, 55.6 % female) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *n* | *M* | *SD* | *ICC* | *n* | *M* | *SD* | *ICC* | *n* | *M* | *SD* | *ICC* |
| Sample characteristics | | | | | | | | | | | | | |
| | Age | 561 | 14.61 | 0.45 | .00 | 718 | 14.63 | 0.47 | .00 | 635 | 14.64 | 0.47 | .01 |
| | Cognitive ability score | 519 | 19.96 | 4.01 | .04 | 681 | 19.99 | 4.22 | .05 | 610 | 19.62 | 4.27 | .01 |
| | Math grade in Grade 8[a] | 557 | 2.81 | 0.97 | .04 | 714 | 2.73 | 0.97 | .06 | 624 | 2.89 | 0.90 | .02 |
| Study variables (subject: mathematics) | | | | | | | | | | | | | |
| T1 | Utility value | 517 | 2.56 | 0.49 | .05 | 680 | 2.52 | 0.47 | .07 | 607 | 2.52 | 0.49 | .09 |
| | Self-concept | 515 | 2.76 | 0.79 | .03 | 678 | 2.74 | 0.81 | .04 | 606 | 2.67 | 0.81 | .05 |
| | Homework self-efficacy | 427 | 2.80 | 0.62 | .03 | 599 | 2.72 | 0.62 | .06 | 514 | 2.71 | 0.65 | .04 |
| | Effort (TR) | 497 | 3.03 | 0.81 | .11 | 695 | 2.92 | 0.85 | .03 | 601 | 2.84 | 0.94 | .05 |
| | Test score | 517 | 48.67 | 16.50 | .08 | 676 | 49.85 | 18.19 | .20 | 600 | 46.26 | 16.75 | .15 |
| T2 | Utility value | 530 | 2.64 | 0.50 | .04 | 680 | 2.53 | 0.51 | .08 | 601 | 2.45 | 0.50 | .08 |
| | Self-concept | 530 | 2.78 | 0.80 | .03 | 679 | 2.73 | 0.80 | .04 | 602 | 2.63 | 0.81 | .06 |
| | Homework self-efficacy | 492 | 2.82 | 0.64 | .05 | 659 | 2.71 | 0.67 | .05 | 586 | 2.67 | 0.65 | .03 |
| | Effort (TR) | 541 | 3.07 | 0.79 | .10 | 719 | 2.89 | 0.90 | .05 | 581 | 2.84 | 0.97 | .05 |
| T3 | Utility value | 516 | 2.60 | 0.49 | .02 | 627 | 2.53 | 0.49 | .11 | 557 | 2.44 | 0.51 | .07 |
| | Self-concept | 514 | 2.84 | 0.76 | .03 | 628 | 2.80 | 0.76 | .05 | 559 | 2.70 | 0.77 | .05 |
| | Homework self-efficacy | 460 | 2.89 | 0.64 | .02 | 581 | 2.82 | 0.65 | .05 | 523 | 2.71 | 0.69 | .04 |
| | Effort (TR) | 540 | 3.05 | 0.85 | .09 | 710 | 2.90 | 0.92 | .07 | 622 | 2.87 | 0.94 | .03 |
| | Test score | 516 | 32.59 | 7.51 | .03 | 634 | 31.85 | 8.55 | .04 | 559 | 30.74 | 8.24 | .02 |

*Notes*. [a] In Germany, the grading system ranges from 1 (best grade) to 6 (worst grade). *n* = number; *M* = mean; *SD* = standard deviation; *ICC* = intraclass correlation coefficient; T = time point; TR = teacher rating.

Table 3

*Intercorrelations of the study variables within class (below diagonal) and between class (above diagonal)*

| | | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T1 | (1) | Utility value | – | .52*** | .58*** | .02 | .23* | .72*** | .55*** | .27* | -.02 | .73*** | .48*** | .41*** | -.03 | .03 |
| | (2) | Self-concept | .38*** | – | .60*** | .10 | .38*** | .29* | .81*** | .38*** | .06 | .35** | .78*** | .46*** | .05 | .14 |
| | (3) | HW Self-efficacy | .31*** | .48*** | – | .06 | .34** | .36** | .60*** | .66*** | .14 | .38*** | .63*** | .54*** | .07 | .09 |
| | (4) | Effort (TR) | .16*** | .28*** | .13*** | – | .28** | .09 | .14 | .03 | .55*** | .02 | .11 | .17 | .50*** | .43*** |
| | (5) | Test score | .23*** | .54*** | .27*** | .42*** | – | .15 | .44*** | .25** | .24* | .12 | .34** | .43*** | .21* | .48*** |
| T2 | (6) | Utility value | .68*** | .29*** | .25*** | .14*** | .19*** | – | .44*** | .29* | .12 | .84*** | .36** | .45*** | .05 | .16 |
| | (7) | Self-concept | .36*** | .84*** | .47*** | .29*** | .54*** | .35*** | – | .44** | .12 | .41*** | .88*** | .55*** | .09 | .29** |
| | (8) | HW Self-efficacy | .30*** | .38*** | .58*** | .15*** | .28*** | .29*** | .46*** | – | .21* | .25$^{\dagger}$ | .46*** | .48** | .19* | .14$^{\dagger}$ |
| | (9) | Effort (TR) | .23*** | .30*** | .16*** | .67*** | .41*** | .21*** | .33*** | .18*** | – | .06 | .07 | .20* | .72*** | .37*** |
| T3 | (10) | Utility value | .60*** | .27*** | .19*** | .11*** | .20*** | .66*** | .30*** | .26*** | .18*** | – | .36*** | .47*** | .12 | .02 |
| | (11) | Self-concept | .32*** | .79*** | .43*** | .25*** | .50*** | .31*** | .85*** | .41*** | .29*** | .35*** | – | .58*** | .04 | .26* |
| | (12) | HW Self-efficacy | .29*** | .41*** | .52*** | .10** | .27*** | .30*** | .48*** | .59*** | .16*** | .33*** | .51*** | – | .19* | .20* |
| | (13) | Effort (TR) | .20*** | .30*** | .15*** | .62*** | .37*** | .18*** | .32*** | .21*** | .71*** | .16*** | .33*** | .17*** | – | .30** |
| | (14) | Test score | .16*** | .43*** | .19*** | .22*** | .51*** | .11*** | .40*** | .20*** | .24*** | .12*** | .40*** | .18*** | .24*** | – |

*Note*. T = time point; HW = homework; TR = teacher rating.

$^{\dagger}$ $p$ < .10; * $p$ < .05; ** $p$ < .01; *** $p$ < .01.

Table 4

*Effects of the relevance interventions on students' competence beliefs, teacher-rated effort, and test-based achievement in mathematics*

| | Self-concept | | | | Homework-related self-efficacy | | | | Teacher-rated effort | | | | Test score | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T2 | | T3 | | T2 | | T3 | | T2 | | T3 | | T3 | |
| | *β* (*SE*) | *p* | *β* (*SE*) | *p* | *β* (*SE*) | *p* | *β* (*SE*) | *p* | *β* (*SE*) | *p* | *β* (*SE*) | *p* | *β* (*SE*) | *p* |
| Student level | | | | | | | | | | | | | | |
| DV at T1 | .84 (.02) | .000 | .79 (.02) | .000 | .59 (.03) | .000 | .53 (.02) | .000 | .67 (.03) | .000 | .63 (.02) | .000 | .55 (.03) | .000 |
| Class level | | | | | | | | | | | | | | |
| DV at T1 | .83 (.06) | .000 | .80 (.06) | .000 | .69 (.08) | .000 | .62 (.09) | .000 | .52 (.10) | .000 | .46 (.09) | .000 | .30 (.06) | .000 |
| Quotations | .10 (.05) | .019 | .09 (.06) | .062 | .16 (.05) | .002 | .20 (.06) | .001 | .14 (.08) | .029 | .12 (.07) | .046 | .18 (.07) | .004 |
| Text | .03 (.05) | .240 | .03 (.05) | .264 | .08 (.06) | .069 | .16 (.06) | .008 | .01 (.07) | .463 | -.01 (.07) | .474 | .06 (.06) | .168 |
| Residuals | | | | | | | | | | | | | | |
| Student level | 0.27 (.01) | .000 | 0.35 (.02) | .000 | 0.63 (.03) | .000 | 0.68 (.03) | .000 | 0.51 (.03) | .000 | 0.58 (.03) | .000 | 0.72 (.03) | .000 |
| Class level | 0.02 (.01) | .000 | 0.02 (.01) | .001 | 0.01 (.01) | .104 | 0.02 (.01) | .057 | 0.05 (.01) | .000 | 0.05 (.01) | .000 | 0.02 (.01) | .002 |

*Note*. Students' gender, pretest cognitive ability score, and end-of-year math grade in Grade 8 were included in the models as auxiliary variables.
*β* = standardized regression coefficient; *SE* = standard error; *p* = one-tailed *p*-value; DV = dependent variable; T = time point.

**Treatment effects at posttest and follow-up**

Treatment effects at posttest and follow-up are reported in Table 4. Concerning mathematics-related self-concept, students in classes in the quotations condition reported statistically significant higher values at the posttest ($\beta$ = .10; $p$ = .019) than students in classes in the control condition, controlling for their initial values. At the follow-up, this effect was slightly smaller and missed statistical significance ($\beta$ = .09; $p$ = .062). The text condition did not show a statistically significant effect on students' mathematics-related self-concept, neither at the posttest ($\beta$ = .03; $p$ = .240) nor at the follow-up ($\beta$ = .03; $p$ = .264).

With regards to homework-related self-efficacy in mathematics, students in classes in the quotations condition reported statistically significant higher values than students in classes in the control condition at both the posttest ($\beta$ = .16; $p$ = .002) and the follow-up ($\beta$ = .20; $p$ = .001). For students in classes in the text condition, no treatment effect on homework-related self-efficacy was observed at the posttest ($\beta$ = .08; $p$ = .069). However, at the follow-up, a statistically significant positive treatment effect emerged ($\beta$ = .16; $p$ = .008), which was not significantly different from the effect of the quotations condition according to a Wald-$\chi^2$-test ($\chi^2$ (1) = 0.37, $p$ = .544).

Concerning students' individual effort in mathematics as rated by their teachers, positive effects of the quotations condition emerged at both the posttest ($\beta$ = .14; $p$ = .029) and the follow-up ($\beta$ = .12; $p$ = .046). The text condition had no statistically significant effect on students' effort as observed by their teachers, neither at the posttest ($\beta$ = .01; $p$ = .463) nor at the follow-up ($\beta$ = -.01; $p$ = .474).

As for mathematics achievement, students in classes in the quotations condition had statistically significant better scores in the speed test ($\beta$ = .18; $p$ = .004) than students in classes in the control condition. Students in classes in the text condition, however, did not perform significantly better at the test ($\beta$ = .06; $p$ = .168) than students in classes in the control group.

## Discussion

What can be done to help secondary school students become more self-confident, work harder, and show higher performance in mathematics? Based on the findings of the present study, a short relevance intervention (90 minutes in class, two booster tasks at home) seems to be a promising support measure. In a cluster randomized controlled experiment, the effectiveness of two relevance interventions including the presentation of examples about the utility of mathematics for various life domains and individual writing assignments differing by condition was compared in mathematics classrooms in Grade 9. Commenting on quotations from peers about the relevance of mathematics fostered students' self-concept, homework-related self-efficacy, teacher-rated effort, and test scores in mathematics until up to five months after the intervention. Writing a text about the relevance of mathematics promoted students' long-term homework-related self-efficacy in mathematics to the same extent as the quotations condition, but no statistically significant effects were found on other outcomes under study.

**New insights into the effectiveness of classroom-based relevance interventions**

Researchers in STEM fields acknowledge a need for relevance-enhanced teaching approaches that are highly effective and implementable by educational practitioners in real-life classroom contexts (e.g., Davis & McPartland, 2012; Osborne & Dillon, 2008). However, experimental studies testing the effectiveness of different relevance interventions under realistic and natural educational conditions are still rare. Using an adequate sample size of 82 ninth-grade classes, the effects of two class-level relevance interventions implemented in a real-life classroom setting on students' competence beliefs, teacher-rated effort, and achievement were assessed in the current study. Such a broad range of important outcomes has rarely been considered in prior motivation intervention studies (see meta-analytic and narrative reviews on motivation interventions in education by Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016). The direct comparison of two treatment conditions, the inclusion of the teacher's perspective, and the use of a follow-up measurement in the treatment evaluation constitute further innovations in classroom-based relevance intervention research.

**The quotations condition: a promising new approach.** The overall pattern of results found in the present study suggests that a newly developed intervention approach including the evaluation of quotations about the relevance of mathematics in young adults' lives was more effective than a strategy adapted from prior research, namely the self-generation of arguments for the relevance of mathematics in a text (e.g., Hulleman & Harackiewicz, 2009). This finding corresponds with the results concerning the effects of the MoMa interventions on students' value beliefs (Gaspard, Dicke, Flunger, Brisson, et al., 2015) and student-rated effort (Gaspard et al., 2016) and may be explained in various ways.

First, although several examples of the utility of mathematics were discussed in the presentation preceding the writing assignment, finding and describing reasons for the relevance of mathematics as a domain in a text might have been a difficult task for the students (see Harackiewicz et al., 2010). Students in the text condition might therefore not have come up with the same number and range of relevance arguments that students read in the quotations.

Second, the writing of a text using reasoned argument—a typical task performed in diverse school subjects—might have been less engaging to students than the comparatively novel task of commenting on quotations. Compared to the text assignment, the novelty of the quotations assignment might thus have resulted in more in-depth and sustained learning about the relevance of mathematics (see Finn & Zimmer, 2012).

Third, differences in the quality of the connections made between mathematical knowledge and students' personal lives might also have contributed to the different pattern of results for the intervention conditions (e.g., Canning & Harackiewicz, 2015; Hulleman & Cordray, 2009; Hulleman et al., 2017). By getting authentic information about the utility of mathematics from older students and young adults that ninth graders can easily connect to, students might have identified with the interviewees and realized that mathematical knowledge will be meaningful to their possible future (e.g., Markus & Nurius, 1986; Oyserman & Destin, 2010). In addition, students in the quotations condition were asked to relate the interviewees' utterances about the utility of mathematics to their personal lives by answering several questions one after the other (see Appendix, Part A). This guided step-by-step procedure might have helped students in the quotations condition to reflect on the personal relevance of mathematics more in depth than students in the text condition (see Acee & Weinstein, 2010, for another example of a successful motivation intervention using a step-by-step guidance to process persuasive messages).

**Promoting students' competence beliefs, effort, and achievement: Are the effects stable?** A closer look at the results of the present study suggests that students' math self-concept was promoted through the quotations condition for six weeks, whereas students' homework-related self-efficacy was fostered through both intervention conditions for five months. These differen-tial treatment effects on students' competence beliefs might pertain to conceptual differences in the nature of these two outcomes (Bong & Skaalvik, 2003): Students' domain-specific self-concept seems to be more stable and less easily malleable than students' homework-related self-efficacy beliefs, which was also reflected in the high predictive power of students' initial math self-concept for students' subsequent math self-concept in the present study (see Table 4). The disappearance of the positive effect of the quotations condition on students' self-concept at the follow-up might have resulted from two processes taking place over time: On the one hand, students may not (yet) have perceived any actual improvement in their math achievement (compared to their previous math achievement or to their performance in other domains). An actual improvement in performance, in turn, has been found to be a precondition of a sustained promotion of students' domain-specific self-concept (see meta-analysis on self-concept interventions by O'Mara, Marsh, Craven, & Debus, 2006). On the other hand, students may have compared their own math achievement with their classmates' math performance. Such internal and external frame of reference processes (Marsh, 1986) could have led to a re-adaption to students' initial levels of math self-concepts over the course of five months.

Another particularly interesting finding is that teachers of classes in the quotations condition rated their students as putting more effort in their math tasks. As effects only occurred in one intervention condition and largely corresponded with findings on students' self-reported effort (Gaspard et al., 2016), it is unlikely that teachers gave a positively biased account of their students' effort due to their awareness of the class' study condition. To the contrary, it could be that the effects found on students' effort were actually underestimated due to the limited objectivity of the teacher ratings. In our sample, 54 % of the teachers had already taught their students in mathematics in previous school years. Additional analyses showed that these teachers' judgments of students' effort were significantly more stable ($r_{T1\text{-}T2}$ = .70) than those of the teachers who had not taught their classes in earlier school years ($r_{T1\text{-}T2}$ = .63; $p$ = .008). Preexisting evaluations of students' attitudes as well as social comparisons between the students in a class, as has been emphasized, for instance, in research on teachers' evaluations of students' achievement (e.g., McMillan, 2001; Südkamp, Kaiser, & Möller, 2012), might then have contributed to an underestimation of the intervention effects on students' effort.

Last but not least, the positive effect of the quotations condition on students' achievement five months after the intervention highlights the potential of this intervention approach in the longer run. The increase in both motivation and effort—factors which are particularly important for students' achievement in standardized test in mathematics (e.g., Cole et al., 2008; Marsh et al., 2005)—might have resulted in the better test performance of students in the quotations condition.

**Limitations and suggestions for future research**

Apart from constraints to the generalizability of the current research findings and the need for replication with other student samples—which applies to all intervention studies—there are four central limitations to the present investigation as well as resulting research suggestions. First, as in Germany students are typically not administered more than one state-based standardized achievement test as used in the pretest within one school year and subject, a different achievement measure had to be used in the posttest. To minimize the risk that students coping better with one of the two types of test in mathematics would be unevenly distributed across control and experimental groups, a huge sample was used and randomization was blocked within school. However, using achievements tests based on the same metrics would have strengthened the study even further.

Second, as this study's focus was on analyzing and comparing the main effects of two relevance interventions, no statements can be made about the mechanisms leading to the differences in the effects on the studied outcomes within and between the intervention conditions. More research is needed to clarify, for instance, why students' math self-concept was promoted only shortly after the intervention whereas students' homework-related self-efficacy was mainly affected five months after the intervention. Similarly, further studies are needed to explore why the quotations condition fostered all of the studied outcomes whereas the text condition only promoted homework-related self-efficacy. Qualitative content analyses of students' writing assignments (e.g., the range and type of relevance arguments found in the text condition, see Canning & Harackiewicz, 2015, studies 2 and 3) and elaborate investigations on students' responsiveness (i.e., the degree to which students worked on the intervention material as intended, e.g., Hulleman & Cordray, 2009), which both are beyond the scope of the current study, might provide additional insights into these open questions. Besides, students' literacy skills might affect the quality of students' writings and, thus, the intervention effects. Investigating the mediating role of students' reading and writing skills in essay-based relevance interventions would be an interesting direction for future research.

Third, the unique contributions of the different elements of the relevance interventions to their effectiveness cannot be disentangled in the present study. Based on theoretical considerations made in EVT (Eccles & Wigfield, 2002) and empirical evidence from prior relevance intervention studies (e.g., Durik, Shechter, et al., 2015; Hulleman & Harackiewicz, 2009; Woolley et al., 2013), three elements were combined: First, a confidence reinforcement was implemented to avoid negative treatment effects on students who believe they cannot improve their math achievement; second, examples about the utility of mathematics were provided to facilitate working on the third element, the individual writing assignments. As students have heterogeneous motivational preconditions and needs, a combination of these different elements was chosen to address a maximum of students. Such a high fit with educational reality is an important prerequisite to enable the scaling up of educational interventions (Cohen & Loewenberg Ball, 2007). It would thus be up to future studies to investigate the importance of the three treatment elements used in the present interventions by creating different conditions with and without these respective elements (e.g., Canning & Harackiewicz, 2015; Durik, Shechter, et al., 2015).

Last but not least, the present interventions have been implemented by trained researchers who were unfamiliar with and to the classes. Future research also needs to examine the effectiveness of the present interventions when mathematics teachers themselves carry them out in their classrooms. When teachers are responsible for implementing an intervention in their classes, there are several sources of infidelity such as the dosage of the intervention or students' responsiveness to the treatment (e.g., Hulleman & Cordray, 2009), which could affect the treatments' effectiveness. Teachers might focus on specific elements of the interventions more strongly than on others or even completely adapt the contents of the treatment based on personal and professional beliefs as well as on their students' motivational features (Cohen & Loewenberg Ball, 2007). Comparing the effectiveness of teacher- and researcher-led relevance interventions with each other would thus be a crucial next step to find the most effective way of implementing the current interventions (cf., implementation science, e.g., Forman et al., 2013).

## Conclusions

Despite its shortness (90 minutes in class, two short booster tasks at home), the present relevance intervention program showed a sustained impact on students' competence beliefs, teacher-rated effort, and test scores in mathematics in a real-life learning setting. Integrating the presentation of utility information and a self-generation task into one approach was particularly impactful when students commented on interview quotations about the utility of mathematics in daily life situations in a writing assignment. The success of this type of relevance intervention in fostering a broad range of important educational outcomes could inspire future researchers to develop further practically relevant and even more sustained motivation interventions in STEM (e.g., by integrating different motivation theories, see Acee & Weinstein, 2010). In addition, the interventions tested in the present study could be extended by including teachers in the implementation process. By taking such further steps, the current investigation could have the potential to contribute to improving educational practice and to attracting more students to STEM-related courses and occupations on a larger scale.

## Footnotes

[1] The ISEI is an international standard measure indicating the status of the occupation, ranging from 16 to 90.

[2] Nine of the teachers taught two classes each.

[3] Unequal class sample sizes in different conditions resulted from the fact that classes whose teachers participated with two classes were deliberately assigned to the same condition. The sample characteristics of each condition can be found in Table 2.

[4] As maximally 1.2 % of the variance in the outcome variables was due to differences between schools, the school level was neglected in the analyses.

## References

Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *The Journal of Experimental Education, 78*(4), 487-512. doi:10.1080/00220970903352753

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 20-31. doi:10.1080/10705511.2014.955104

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71-81). New York: Academic Press.

Baumert, J., Gruehn, S., Heyn, S., Köller, O., & Schnabel, K. U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation – Band 1: Skalen Längsschnitt Welle 1-4 [Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU). Documentation – Volume 1: Scales of Data Collections 1-4]*. Berlin: Max-Plank-Institut für Bildungsforschung.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40. doi:10.1023/A:1021302408382

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility value information. *Motivation Science, 1*(1), 47-71. doi:10.1037/mot0000015

Chouinard, R., Karsenti, T., & Roy, N. (2007). Relations among competence beliefs, utility value, achievement goals, and effort in mathematics. *British Journal of Educational Psychology, 77*, 501-517. doi:10.1348/000709906x133589

Chouinard, R., & Roy, N. (2008). Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *British Journal of Educational Psychology, 78*, 31-50. doi:10.1348/000709907x197993

Cohen, D. K., & Loewenberg Ball, D. (2007). Educational innovation and the problem of scale. In B. Schneider & S.-K. McDonald (Eds.), *Scale-up in education: Ideas in principle (Vol. 1)* (pp. 19-36). Plymouth, UK: Rowman & Littlefield Publishers.

Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology, 33*(4), 609-624. doi:10.1016/j.cedpsych.2007.10.002

Craven, R. G., Marsh, H. W., Debus, R. L., & Jayasinghe, U. (2001). Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. Journal of Educational Psychology, 93(3), 639-645. doi:10.1037/0022-0663.93.3.639

Davis, M. H., & McPartland, J. M. (2012). High school reform and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 515-539). New York: Springer.

Denissen, J. J., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430-447. doi:10.1111/j.1467-8624.2007.01007.x

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*(3), 597-610. doi:10.1037/0022-0663.99.3.597

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

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 interest and performance. *Motivation and Emotion, 39*, 104-118. doi:10.1007/s11031-014-9419-0

Eccles, J. S. (2005). Subjective task values and the Eccles et al. model of achievement related choices. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105-121). New York: Guilford.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motives: *Psychological and sociological approaches* (pp. 75-146). New York: Freeman.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225. doi:10.1177/0146167295213003

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132. doi:10.1146/annurev.psych.53.100901.135153

Enders, C. K. (2010). Applied missing data analysis. New York: Guilford Press.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121-138. doi:10.1037/1082-989x.12.2.121

Ennemoser, M., Krajewski, K., & Schmidt, S. (2011). Entwicklung und Bedeutung von Mengen-Zahlen-Kompetenzen und eines basalen Konventions-und Regelwissens in den Klassen 5 bis 9 [Development and importance of sets and numbers competencies and of a basic knowledge of conventions and rules in classes 5 through 9]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie / German Journal of Developmental and Educational Psychology, 43*(4), 228-242. doi:10.1026/0049-8637/a000055

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97-132). New York: Springer.

Forman, S. G., Shapiro, E. S., Codding, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., . . . Stoiber, K. C. (2013). Implementation science and school psychology. *School Psychology Quarterly, 28*(2), 77-100. doi:10.1037/spq0000019

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109. doi:10.3102/00346543074001059

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*(9), 1226-1240. doi:10.1037/dev0000028

Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B. M., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in mathematics classrooms on motivation in verbal domains. *AERA Open, 2*(2), 1-14. doi:10.1177 /2332858416649168

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663-677. doi:10.1037/edu0000003

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*(1), 549-576. doi:10.1146/annurev.psych.58.110405.085530

Haracklewicz, J. M., Hulleman, C. S., Rozek, C. S., Katz-Wise, S., & Hyde, J. S. (2010, March). *Parents' understanding of the utility value of STEM courses for high school students*. Paper presented at the 2010 Biennial Meeting of the Society for Research on Adolescence, Philadelphia, PA.

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological Science, 23*(8), 899-906. doi:10.1177/0956797611435530

Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klasse [Cognitive Ability Test for Grades 4 through 12]. Göttingen, Germany: Hogrefe.

Howell, D. C. (2012). *Statistical methods for psychology (8th ed.).* Belmont, CA: Cengage Learning.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*(1), 88-110. doi:10.1080/19345740802539325

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*(4), 880-895. doi:10.1037/A0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412. doi:10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*(3), 387-404. doi:10.1037/edu0000146

Husman, J., & Hilpert, J. (2007). The intersection of students' perceptions of instrumentality, self-efficacy, and goal orientations in an online mathematics course. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology, 21*(3-4), 229-239. doi:10.1024/1010-0652.21.3.229

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*(2), 509-527. doi:10.1111/1467-8624.00421

Korendijk, E. J. H., Hox, J. J., Moerbeek, M., & Maas, C. J. M. (2011). Robustness of parameter and standard error estimates against ignoring a contextual effect of a subject-level covariate in cluster-randomized trials. *Behavior Research Methods, 43*(4), 1003-1013. doi:10.3758/s13428-011-0094-8

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research, 86*(2), 602-640. doi:10.3102/0034654315617832

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*(9), 954-969. doi:10.1037/0003-066X.41.9.954

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129-149. doi:10.3102 /00028312023001129

Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology, 19*(3), 119-129. doi:10.1024/1010-0652.19.3.119

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B. O., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764-802. doi:10.1080/00273170903333665

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397-416. doi:10.2307/3696511

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20-32. doi:10.1111/j.1745-3992.2001.tb00055.x

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide, seventh edition*. Los Angeles, CA: Muthén & Muthén.

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181-206. doi:10.1207/s15326985ep4103_4

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*. London: The Nuffield Foundation.

Oyserman, D., & Destin, M. (2010). Identity-based motivation: Implications for intervention. The *Counseling Psychologist, 38*(7), 1001-1043. doi:10.1177/0011000010374775

Peetsma, T., & van der Veen, I. (2011). Relations between the development of Future Time Perspective in three life domains, investment in learning, and academic achievement. *Learning and Instruction, 21*(3), 481-494. doi:10.1016/j.learninstruc.2010.08.001

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, R., Neubrand, M., . . . Schiefele, U. (2006). *Pisa 2003: Dokumentation der Erhebungsinstrumente [Pisa 2003: Documentation of assessment instruments]*. Münster: Waxmann Verlag.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*(2), 173-185. doi:10.1037/1082-989x.2.2.173

Reschly, A. L., & Christenson, S., L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3-20). New York: Springer.

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *Elementary School Journal, 100*(5), 443-471. doi:10.1086/499650

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM Motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist, 51*(2), 146-163. doi:10.1080 /00461520.2016.1154792

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. doi:10.1037110003-066X.55.1.68

Schmidt, S., Ennemoser, M., & Krajewski, K. (Eds.). (2013). *Deutscher Mathematiktest für 9. Klassen [German mathematics test for Grade 9]*. Göttingen: Hogrefe.

Schuitema, J., Peetsma, T., & van der Veen, I. (2014). Enhancing student motivation: A longitudinal intervention study based on Future Time Perspective theory. *The Journal of Educational Research, 107*(6), 467-481. doi:10.1080/00220671.2013.836467

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*(3), 303-317. doi:10.1177/0146167210396380

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences (5th ed.)*. London: Routledge.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*(3), 743-762. doi:10.1037/a0027627

Trautwein, U., & Lüdtke, O. (2009). Predicting homework motivation and homework effort in six school subjects: The role of person and family characteristics, classroom factors, and school track. *Learning and Instruction, 19*(3), 243-258. doi:10.1016/j.learninstruc.2008.05.001

Trautwein, U., Lüdtke, O., Kastens, C., & Köller, O. (2006). Effort on homework in grades 5-9: Development, motivational antecedents, and the association with effort on classwork. *Child Development, 77*(4), 1094-1111. doi:10.1111/j.1467-8624.2006.00921.x

Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97*(6), 1115-1128. doi:10.1037/a0017048

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*(2), 438-456. doi:10.1037/0022-0663.98.2.438

Trautwein, U., Nagengast, B., Marsh, H. W., Gaspard, H., Dicke, A.-L., Lüdtke, O., & Jonkmann, K. (2013). Expectancy-value theory revisited: From expectancy-value theory to expectancy-valueS theory? In D. M. McInerney, H. W. Marsh, & R. G. Craven (Eds.), *Theory driving research: New wave perspectives on self-processed and human development* (pp. 233-249). Charlotte, NC: Information Age Publishing.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55-66). London: National Foundation for Educational Research.

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science, 23*(1), 73-82. doi:10.1177/0963721413512856

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1-35. doi:10.1016/J.Dr.2009.12.001

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg (Ed.), *Handbook of child psychology (6th ed., vol. 3)* (pp. 933-1002). Hoboken, NJ: John Wiley & Sons.

Woolley, M. E., Rose, R. A., Orthner, D. K., Akos, P. T., & Jones-Sanpei, H. (2013). Advancing academic achievement through career relevance in the middle grades: A longitudinal evaluation of CareerStart. *American Educational Research Journal, 50*(6), 1309-1335. doi:10.3102/0002831213488818

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267-301. doi:10.3102/0034654311405999

Zimmerman, B. J., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology, 30*(4), 397-417. doi:10.1016/j.cedpsych.2005.05.003

# 6

# Study 3

**Who sticks to the instructions—and does it matter?**

**Antecedents and effects of students' fidelity to a classroom-based relevance intervention**

Brisson, B. M., Hulleman, C. S., Häfner, I., Gaspard, H., Flunger, B., Dicke, A.-L., Trautwein, U., & Nagengast, B., (revise and resubmit). Who sticks to the instructions—and does it matter? Antecedents and effects of students' fidelity to a classroom-based relevance intervention. *Journal of Research on Educational Effectiveness*.

**Abstract**

Why do some students benefit from interventions and others do not? By investigating the antecedents and effects of students' fidelity to a classroom-based relevance intervention, the current study aims to shed a closer light on the intervention processes that make relevance interventions in real-life educational settings work—or not work. Using data from a cluster-randomized controlled experiment with 1916 ninth-grade students, intervention fidelity was assessed as students' responsiveness to two written intervention activities regarding the personal relevance of mathematics (evaluating quotations or writing a text). Based on the hypothesized theory of change for relevance interventions, 1280 student essays completed during the intervention were coded on three indicators of responsiveness (positive arguments, personal connections, in-depth reflections) which were combined into a continuous responsiveness index. Linear regression analyses showed that students' conscientiousness, gender, math-related motivation and achievement predicted the responsiveness index. Complier-average causal effects analyses revealed that intervention effects on the target outcome, students' utility value beliefs, compared to the control group, were stronger for responsive than for nonresponsive students, particularly in the text condition. This research highlights the importance of investigating intervention processes in order to optimize the theories and designs of classroom interventions.

**Keywords:** intervention fidelity · student responsiveness · utility-value beliefs · relevance intervention · classroom intervention · motivation · complier-average causal effects

## Introduction

Relevance interventions aim at improving academic outcomes by raising students' beliefs about the utility value of a task or subject (e.g., Durik, Hulleman, & Harackiewicz, 2015; Eccles et al., 1983). In real-life classroom settings, relevance interventions have indeed been shown to boost the target outcome, students' utility value beliefs, which can mediate effects on more distal outcomes such as interest, effort, and achievement in science-related subjects (Rosenzweig & Wigfield, 2016). Yet, little is still known about the intervention processes that trigger a change in students' utility value beliefs themselves. Studying students' responsiveness to the intervention as one important element of intervention fidelity (i.e., the question of whether the intervention as implemented corresponds with the intervention as designed) is one way to gain insights into the process through which the target outcome is affected (O'Donnell, 2008). In addition, fidelity data can be useful for more precisely estimating the effectiveness of educational interventions and for finding explanations why the same intervention works in one context but not in another (e.g., Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Relevance interventions are typically based on individual writing tasks that are supposed to trigger students' reflection about relevance and to enable the personalization of the intervention message—processes which are assumed to cause a change in students' motivational beliefs (e.g., Lazowski & Hulleman, 2016; Yeager & Walton, 2011). In classroom-based experiments, researchers may deliver educational interventions as intended, but the students may not do what they are expected to do. In turn, if students do not stick to the instructions of the intervention task, the processes initiating a change in students' beliefs may not unfold (i.e., intervention processes; Murrah, Kosovich, & Hulleman, 2017). Assessing the core elements of students' responsiveness—the extent to which students complete the intervention activities as intended (Dane & Schneider, 1998)—and analyzing their effect on the target psychological process (i.e., changes in utility value beliefs) might therefore shed light on the process of how relevance interventions work (e.g., Nelson, Cordray, Hulleman, Darrow, & Sommer, 2012)—and why they sometimes do not work (e.g., Husman, Nelson, & Cheng, 2017; Karabenick, Albrecht, & Rausch, 2017).

Initial descriptive studies indicate that the quality of students' responsiveness to writing tasks used in relevance interventions impacts the strength of the intervention effects and is related to subgroup effects (e.g., Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016; Hulleman & Cordray, 2009). However, research is missing into which students' responsiveness to classroom-based relevance interventions is systematically defined, assessed, and analyzed as the core intervention elements leading to changes in students' motivational beliefs. Using both descriptive and causal approaches, the current study aims to fill this gap in research by (a) providing a theory of change to assess and combine indicators of students' responsiveness to a math-related relevance intervention, (b) exploring individual student characteristics and classroom perceptions as predictors of a combined index of students' responsiveness, and (c) examining to what extent different degrees of responsiveness affected the target psychological process of the intervention, students' math-related utility value beliefs.

**Effectiveness of relevance interventions**

Based in expectancy-value theory (Eccles et al., 1983; Eccles & Wigfield, 2002), students' utility value beliefs (i.e., the perceived usefulness of an academic activity or domain to a student's current or future goals) are the target *psychological process* of relevance or utility-value interventions, its proximal outcome. Numerous correlational studies have shown that students who perceive the learning contents as highly useful feature high levels of other value beliefs (e.g., intrinsic and attainment values), competence beliefs (e.g., self-efficacy and ability perceptions), and effort (for overviews, see e.g., Roeser, Eccles, & Sameroff, 2000; Wigfield & Cambria, 2010; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006; for correlations in the current sample, see Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson, et al., 2015). Utility value beliefs were also found to predict achievement longitudinally (e.g., Hulleman, Durik, Schweigert, & Harackiewicz, 2008). Accordingly, fostering students' utility value beliefs in targeted interventions may lead to positive effects on other important academic outcomes.

Like other social-psychological interventions in education (cf., Yeager & Walton, 2011), relevance interventions often rely on the assumption that a change in students' personal beliefs can be caused through individual writing exercises (Lazowski & Hulleman, 2016). Such writing tasks may require students, for example, to write an essay about the personal relevance of a self-chosen topic from their science class (e.g., Hulleman & Harackiewicz, 2009). The relevance intervention program "Motivation in Mathematics" (MoMa, Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson et al., 2015) utilized in the current study consisted of a 90-minute researcher-led session in ninth-grade math classes. During the intervention session, students first watched a presentation on examples about the relevance of mathematical skills for various areas of life and then completed an individual writing task. Students either commented on the personal relevance of interview statements about situations in which young adults needed mathematical skills (quotations condition) or freely wrote a text about the personal relevance of mathematics (text condition).

The effectiveness of relevance interventions is typically estimated in comparison to a control group of students who either performed an unrelated writing task (e.g., Hulleman & Harackiewicz, 2009) or who did not do any assignment—as was the case in the MoMa study. Overall, relevance interventions using writing exercises have been shown to promote students' utility value beliefs as the target psychological process as well as more distal outcomes including students' competence beliefs, interest, effort, and achievement in subjects like mathematics, physics, biology, and psychology (Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson et al., 2015; Harackiewicz et al., 2016; Hulleman, Godes, Hendricks, & Harackiewicz, 2010, study 2; Hulleman & Harackiewicz, 2009; Hulleman, Kosovich, Barron, & Daniel, 2017, study 2). As assumed, some of the intervention effects—for example, on interest and achievement—were mediated through an increase in utility value beliefs (e.g., Hulleman et al., 2010). Besides, intervention effects have been shown to vary depending on individual student characteristics: Less confident students, low achievers, and first-generation ethnic minority students were found to benefit most from classroom-based relevance interventions (Harackiewicz et al., 2016; Hulleman et al., 2010, study 2; Hulleman & Harackiewicz, 2009; Hulleman et al., 2017, study 2).

The MoMa relevance interventions successfully fostered students' utility value beliefs as well as students' intrinsic and attainment value beliefs, academic self-concept, homework-related self-efficacy, teacher-rated effort, and achievement in mathematics until up to five months after the intervention (Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson et al., 2015). The quotations condition had stronger effects on all outcomes than the text condition. Furthermore, the text condition promoted girls' value beliefs more than boys' (Gaspard, Dicke, Flunger, Brisson et al., 2015).

The success of relevance interventions to boost important academic outcomes has been summarized in meta-analytic and narrative reviews of motivation interventions (Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016; Yeager & Walton, 2011). Yet, not much is known about the intervention processes through which a change in the target psychological process, students' beliefs about the utility of a task or subject, is triggered. Experimental studies have revealed links between the length and contents of students' essays written during the intervention and intervention effects on students' utility value beliefs (Harackiewicz et al., 2016; Hulleman et al., 2017; Hulleman & Cordray, 2009). However, to identify the intervention processes that contribute to the effectiveness of classroom-based relevance interventions, the role of students' responsiveness to the intervention task in manipulating students' utility value beliefs must be studied more systematically (cf., Nelson et al., 2012).

### Student responsiveness—a central but neglected element of intervention fidelity

Research on intervention fidelity is concerned with whether the intervention as received corresponds with the intervention as intended (e.g., Nelson et al., 2012). In classroom-based experiments, variation in intervention fidelity may be attributable to the implementers (e.g., teachers) and, for example, their adherence to the intervention manual, but also to the participants (e.g., students) and their responsiveness, that is, the extent to which they are engaged in the intervention activities (Dane & Schneider, 1998). Such responses to the intervention from both implementers and participants can be classified as *intervention processes* (Murrah et al., 2017).

Establishing, assessing, and reporting fidelity criteria and associated intervention processes in experimental research is important for numerous reasons. It is necessary to make interventions replicable in a standardized way and it supports the internal validity of the study by verifying a baseline level of implementation and enabling a more precise estimation of the intervention effects (Hohmann & Shear, 2002; Mowbray, Holter, Teague, & Bybee, 2003; cf., results by Nagengast et al., 2018). In addition, ensuring high intervention fidelity is greatly desirable as studies with higher levels of fidelity produce stronger intervention effects than those with implementation problems (see reviews and meta-analyses on fidelity in educational interventions by Dane & Schneider, 1998; Durlak et al., 2011; Durlak & DuPre, 2008; O'Donnell, 2008). In particular, studies on intervention fidelity can help to provide explanations for how an intervention works or why it only works for certain subgroups of students (Murrah et al., 2017).

Despite its importance, assessing intervention fidelity has often been neglected in educational experiments—in particular, on students' responsiveness to the intervention activities (Durlak & DuPre, 2008; O'Donnell, 2008). This is an important shortcoming in classroom-based experiments, in which individual writing assignments for students are often used to trigger the target psychological process, a change in students' personal beliefs (e.g., Murrah et al., 2017; Yeager & Walton, 2011). In relevance interventions, making students stick to the instructions of the writing activity, which have been formulated in a way to initiate the manipulation of students' perceived utility value, seems crucial for the effectiveness of the intervention (cf., Lazowski & Hulleman, 2016). At the same time, in a classroom context, it is difficult to control students' responsiveness to the intervention task: While the writing activity can easily be implemented in a highly standardized way (e.g., through trained researchers), it is difficult to ensure that all students complete the individual activity as intended, which, in turn, may affect the effectiveness of the intervention.

**Student responsiveness in relevance interventions**

**Core elements of student responsiveness.** In order to assess students' responsiveness to activities of relevance interventions, the theoretical elements which are assumed to cause a change through the intervention need to be identified in the *theory of change* or *change model* (Nelson et al., 2012). In an intervention theory of change, interventions are expected to instigate two types of processes. The first are intervention processes, which are the actions of both implementers and participants. The second are psychological processes, which are the proximal outcomes of the intervention. These processes work in a chain—intervention processes induce psychological processes which instigate intervention outcomes (cf., Murrah et al., 2017). As previously reviewed, utility value beliefs represent the target psychological process that leads to further intervention effects: These value beliefs lead students to become more engaged in learning, exert more effort, become more interested, and attain better learning outcomes. As the psychological processes of relevance interventions have been established in prior studies (e.g., Hulleman et al., 2010), the focus of the current study is on the intervention processes.

In terms of intervention processes, asking students to write about the relevance of a learning matter is typically the most basic instruction of individual tasks used in relevance interventions (e.g., Durik et al., 2015). *Using relevance arguments* is thus assumed to be the first and most important core element students have to respond to so that the writing task will have a positive effect on their value beliefs. In addition, as utility value beliefs are subjective (i.e., they are high if students consider academic tasks or subjects as useful for personal, not impersonal, goals; Eccles et al., 1983), students are instructed to connect the relevance arguments to their personal lives. *Making personal connections* thus constitutes the second core element of students' responsiveness to the writing activity (see also Hulleman & Cordray, 2009; Hulleman et al., 2017). Furthermore, relevance interventions are assumed to be effective when the students are strongly cognitively engaged in the writing activity (Harackiewicz et al., 2016). From a constructivist learning perspective, sustained learning is enabled when students individually discover information and transform it in a way that it matches with their prior knowledge (e.g., Cunningham & Duffy, 1996; Mayer, 2002, 2004). Accordingly, students who reformulate relevance information provided during an intervention in their own words or who transfer it to personally important contexts in the writing activity may learn about relevance in a more sustained way than students who merely reproduce previously encountered relevance arguments without much reflection. *Using in-depth reflections* may thus represent the third core element of student responsiveness.

**Empirical studies on students' responsiveness.** Hulleman and Cordray (2009) provided some initial evidence of the importance of students' responsiveness to classroom-based relevance intervention tasks. The authors investigated why a relevance intervention in high-school classrooms (writing about the personal relevance of a topic from science class) was less effective than a similar intervention in the laboratory. Analyses on the content of students' essays produced during the intervention showed that students in the classroom failed to make high quality, personal connections to the learning material in their essays, resulting in decreased relative intervention strength compared to the laboratory experiment.

In a second study, Harackiewicz et al. (2016) examined students' responsiveness within an online relevance intervention at university (writing about the personal relevance of a concept from biology class), which was found to foster course achievement for first-generation underrepresented minority (FG-URM) students. As FG-URM students wrote longer essays and used more words indicative of social processes and of cognitive involvement than students not belonging to this group, the authors concluded that these aspects contributed to the success of the intervention for FG-URM students.

Going beyond descriptive analyses, Nagengast et al. (2018) used the measure of student responsiveness developed in the current paper to compare the effects of the MoMa interventions obtained from complier-average causal effects (CACE) models (which take into account students' responsiveness status; e.g., Sagarin et al., 2014) with those obtained from intent-to-treat (ITT) analyses (which are only based on students' group assignment, i.e., experimental vs. control group; e.g., Boruch, 1997). Using various outcomes like students' math-related motivational beliefs and achievement, the authors not only found the estimates obtained from CACE models to be greater than those calculated with ITT analyses but also detected further effects when looking at responsive and nonresponsive students separately. Interestingly, the CACE estimates were found to differ more from the ITT estimates in the text condition than in the quotations condition, hinting at a higher importance of the responsiveness measure in the text condition than in the quotations condition. Students' responsiveness was also found to partially explain the differential effects favoring girls over boys (Gaspard, Dicke, Flunger, Brisson et al., 2015).

**Research gaps.** Although these initial studies have provided important information about links between intervention and psychological processes, further research is warranted for several reasons. First, following recommendations made by Nelson et al. (2012), a clear theory of change is needed to guide the assessment of the core elements of students' responsiveness to relevance interventions and their combination into an index. Such a framework explicates the intervention processes and describes the pathway through which they instigate psychological processes.

Second, it is still unclear which students stick to the instructions of the writing activities used in relevance interventions—and which ones do not. The study by Harackiewicz et al. (2016) hints at a potential link between the characteristics of highly responsive students and those of the students benefitting most from relevance interventions. Identifying individual characteristics predicting students' responsiveness more comprehensively may be helpful to explain previously found differential intervention effects or to hint at further possible moderators of an intervention. In turn, adapting the intervention material to the needs of less responsive students might further increase intervention effects (e.g., Nelson et al., 2012).

Third, to learn more about the intervention processes initiating a change in students' motivational beliefs, studies are needed that go beyond descriptive analyses as conducted by Hulleman and Cordray (2009) or Harackiewicz et al. (2016). CACE models (e.g., Sagarin et al., 2014) are a valuable tool to investigating intervention effects for responsive students (i.e., compliers) and nonresponsive students (i.e., noncompliers) separately (cf., study by Nagengast et al., 2018). However, when intervention responsiveness is assessed through a continuous index as recommended by Nelson et al. (2012), it is difficult to determine a minimal level of responsiveness for defining "compliant" students. Using a random minimal level of responsiveness to group students into compliers and noncompliers as done by Nagengast et al. (2018), however, results in a black-and-white picture which does not provide any information on the dose-response relationship between the *degree* of students' responsiveness and the strength of the intervention effects. Including descriptive information on the core elements of responsiveness and on the continuous index, as well as calculating CACE analyses using different minimal levels of responsiveness for defining a "compliant" student is needed to better understand which elements and degrees of students' responsiveness matter most for changing students' utility value beliefs.

### Potential predictors of student responsiveness

**Stable student characteristics and domain-specific motivation.** Several research studies found secondary school students' cognitive abilities, conscientiousness, and domain-specific motivation (e.g., self-concept, homework-related self-efficacy, and value beliefs related to mathematics) to be positively associated with homework compliance in diverse school subjects (Trautwein & Lüdtke, 2007, 2009; Trautwein, Lüdtke, Schnyder, & Niggli, 2006). Because the writing tasks of the MoMa study were completed in a similar manner as typical homework tasks—which are guided but not necessarily controlled by the teacher (e.g., Cooper, 1989)—these variables might also predict students' responsiveness to written intervention activities. In addition, as girls seem to be particularly compliant with homework in language-related subjects (Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009), which often includes text production tasks, girls might also respond better to written intervention activities than boys.

**Academic achievement.** Numerous studies have found high positive correlations between students' academic achievement and their task engagement (for reviews, see e.g., Fredricks, Blumenfeld, & Paris, 2004; Reschly & Christenson, 2012). Although achievement is sometimes considered an outcome of high task engagement, the relationship is probably reciprocal: Students tend to engage in subjects they are good at (e.g., Eccles & Wigfield, 2002; Finn & Zimmer, 2012). This is why particularly high achievers might get involved in writing activities completed during relevance interventions and thus feature high levels of responsiveness (cf., Harackiewicz et al., 2016, on positive links between students' achievement and the number of words and of personal connections in their relevance essays).

**Classroom perceptions.** One important influence on students' academic engagement in the classroom are peers—especially in teenage years (for reviews, see e.g., Fredricks et al., 2004; Juvonen, Espinoza, & Knifsend, 2012). More precisely, perceived classmates' math-related value beliefs has been found to positively correlate with students' interest, value beliefs, and positive emotions in math class (Frenzel, Goetz, Pekrun, & Watt, 2010; Frenzel, Pekrun, & Goetz, 2007; Schreier et al., 2014). Furthermore, a good classroom structure as indicated by a high disciplinary climate is known to positively correlate with time on task and task involvement (for reviews, see e.g., Fredricks et al., 2004; Reschly & Christenson, 2012). If students perceive their classmates to highly value a subject and the disciplinary climate to be high—in short, the classroom atmosphere to be good—they might also be ready to work on in-class intervention tasks in a concentrated and thorough way.

### The present study

Using data from the MoMa project, the current study presents a theoretical framework to assess the degree of students' responsiveness to writing tasks completed during an intervention about the relevance of mathematics (writing a text or evaluating quotations). By investigating the antecedents and effects of student responsiveness insights into the mechanisms contributing to differences in the effectiveness of the two MoMa intervention approaches are to be gained. Using descriptive and causal methods to explore students' responsiveness as part of a generalizable process triggering a change in students' motivational beliefs, this research can serve as a blueprint for in-depth investigations on the role of the fidelity of implementation for the effectiveness of classroom experiments (e.g., Shnabel, Purdie-Vaughns, Cook, Garcia, & Cohen, 2013).

Based on the assumed theory of change (cf., Nelson et al., 2012; see Method section), students' essays written during the 90-minute MoMa relevance intervention in the classroom were coded on three core elements of students' responsiveness: relevance arguments, personal connections, and in-depth reflections, which were combined into a continuous index. In contrast to prior fidelity research on relevance interventions (Harackiewicz et al., 2016; Hulleman et al., 2017; Hulleman & Cordray, 2009), the assessment of students' responsiveness in the current study was guided by a clear theoretical rationale. By testing whether the hypothesized intervention processes were instantiated and whether they make a causal impact on psychological processes, the current study is aimed at revealing how relevance interventions work (cf., Murrah et al., 2017). Taking into account different degrees of responsiveness, the current study also provides hitherto unknown insights into what happens at the "edges"—with the most and the least responsive students—after the relevance intervention, and into what are the characteristics of these students.

Three research questions were addressed. First, how did students comply with the core elements of the writing tasks completed during the MoMa relevance interventions? Second, which individual student characteristics and classroom perceptions predicted students' responsiveness to the writing tasks? Third, how does the degree of students' responsiveness relate to the effects of the interventions on students' math-related utility value beliefs six weeks and five months after the intervention?

Based on prior research (e.g., Trautwein & Lüdtke, 2009), we expected students' gender, cognitive ability, conscientiousness, math-related achievement and motivation (self-concept, homework-related self-efficacy, intrinsic value, utility value), and classroom perceptions (classmates' math-related value beliefs, disruptions in math class) to be associated with students' intervention responsiveness. Drawing on research on intervention fidelity and the effectiveness of educational experiments in general (e.g., Durlak et al., 2011) and on results by Hulleman and Cordray (2009) and Nagengast et al. (2018) in particular, we furthermore assumed intervention effects on students' math-related utility value beliefs compared to the control group to be strongest for the most responsive students.

## Method

### Sample and procedure

Data were collected as part of the cluster-randomized field experiment "Motivation in Mathematics" (MoMa, Gaspard, Dicke, Flunger, Brisson, et al., 2015) with 82 ninth-grade classes in 25 German academic track schools ("Gymnasium"). A total of 1978 students with active parental consent participated in the study (participation rate: 96.0 %). Sixty-two students absent during the intervention were excluded from the current study, yielding a total sample of 1916 students ($M_{age}$ = 14.62, $SD$ = 0.47; 53.5 % female). Instead of treating these students as non-responsive to the interventions (cf., Nagengast et al., 2018), excluding them from the analyses enabled us to take a closer look at those nonresponsive students who did complete the relevance task but not as expected.

In the beginning of the study, teachers and their classes were randomly assigned within each school to either one of two intervention conditions, “quotations” (25 classes, 561 students; 52.8 % female) or “text” (30 classes, 720 students; 52.4 % female), or to the waiting control group (27 classes, 635 students; 55.6 % female). Afterwards, students took part in three data collections from September 2012 to March 2013: Students in the experimental conditions completed questionnaires before the intervention (T1), as well as six weeks (T2) and five months (T3) after the intervention. Students in the waiting control group completed the same questionnaires at the same time points but did not receive any intervention before the last data collection. All 82 classes fully completed all waves of data collections.

Data on students’ responsiveness were obtained by coding a total of 1280 essays produced during the interventions in class (1281 students wrote relevance essays, but one of them got lost). Data on students’ gender and mathematics achievement at the beginning of the school year were provided by the teachers. Students’ cognitive abilities, conscientiousness, math-related motivation, and classroom perceptions were measured at T1. To analyze the intervention effects, students’ math-related utility value was measured at T1, T2, and T3.

**The MoMa relevance interventions**

After the first data collection, students in both experimental groups took part in a 90-minute standardized intervention about the relevance of mathematics led by trained researchers. The interventions started with a psychoeducational presentation, of which the first part served to reinforce students’ competence beliefs (cf., Canning & Harackiewicz, 2015; Durik et al., 2015). In the second part, students were provided with various examples of the utility of mathematics for future education, career opportunities, and leisure time activities. Right after the presentation, students completed an individual writing assignment differing by condition. Based on theories postulating that students can learn from persons they identify with (Bandura, 1977; Markus & Nurius, 1986; Oyserman & Destin, 2010), students in the quotations condition commented on the personal relevance of six interview statements from young adults describing everyday situations in which they needed mathematical skills. Adapted from Hulleman et al. (2009; 2010), students in the text condition were asked to collect arguments for the personal relevance of mathematics to their current and future lives, and to then write a coherent text detailing their notes (see Appendix, Part A).

Researchers recorded the actual procedure of each intervention in the minutes. Overall, the implementation of the interventions seemed highly standardized: All students in the experimental groups had the occasion to complete the writing assignment during the in-class intervention session. Deviations from the standard procedure occurred in only two out of 55 classes (both in the text condition): In one class, the initial presentation had to be held without any projector due to technical problems; in the other class, students did not work quietly on their individual writing tasks.

Researchers collected students' handwritten essays at the end of the intervention. In addition, students received a portfolio including two intervention boosters to be filled out at home after 1 week and 2 weeks, respectively. In the first booster, students were asked to briefly summarize what they remembered from their individual writing task completed in class during the intervention. The second booster differed by condition and corresponded to the type of intervention task completed in class (quotations: evaluating given relevance information; text: self-generating relevance arguments). More information on students' booster tasks, which were also coded and analyzed, can be found in the Appendix, Part F.

Students in classes in the waiting control condition did not watch any presentation or complete any individual writing tasks. However, they received the on average more powerful intervention (i.e., evaluating quotations, see Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson, et al., 2015) after the last measurement point.

**Assessing student responsiveness**

**Nelson et al.'s (2012) model of intervention fidelity.** To produce the responsiveness data, we followed a systematic model by Nelson et al. (2012) suggesting a number of guidelines to identify and measure fidelity indicators in educational experiments. After defining the core elements of the *change model* (as introduced earlier), the *operational model* specifies how these elements are operationalized in the intervention activities (e.g., by writing an essay with a certain content). Fidelity data are then produced by analyzing the extent to which the elements of the operational model have been executed as intended in the products of the intervention (e.g., students' essays). After determining the reliability and validity of the fidelity data, various fidelity indicators are combined into a fidelity index by choosing a method which is in line with the theory of the intervention. Typically, a fidelity scale is developed by summing across indicators or by giving different weights to the indicators depending on their assumed theoretical importance in affecting the intervention outcome. As a last step, the fidelity index can be related to the outcomes of the intervention.

**Intervention models.** The change model and the operational model of the MoMa interventions are presented in Figure 1. As shown in the change model, we specified the key intervention processes as (a) describing arguments about the usefulness of mathematics (i.e., relevance arguments), which (b) relate to the individual (i.e., contain personal connections) and which (c) represent in-depth reflections. As dealing with the relevance of the learning matter was considered the most essential element of relevance interventions (cf., Durik et al., 2015), this element was assumed to take on a more important role than the other two components in the theory of change (i.e., as a prerequisite for initiating the desired change).

The operational model (see Figure 1) was developed assuming that the intervention components were realized in the use of certain key words or types of words in students' essays (see e.g., Harackiewicz et al., 2016; Pennebaker, Mehl, & Niederhoffer, 2003). Relevance arguments are reflected in words such as "useful", "relevant", "important", and the like. Personal connections materialize in the use of self-references (i.e., first person pronouns, Tausczik & Pennebaker, 2010). A high degree of reflection is represented by describing relevance arguments that go beyond the ones presented in the initial part of the intervention. In line with the theory of change, we assumed that only if students use words indicative of the usefulness of mathematics during the intervention activity, the intervention can trigger an increase in the target psychological process, students' math-related utility value beliefs.

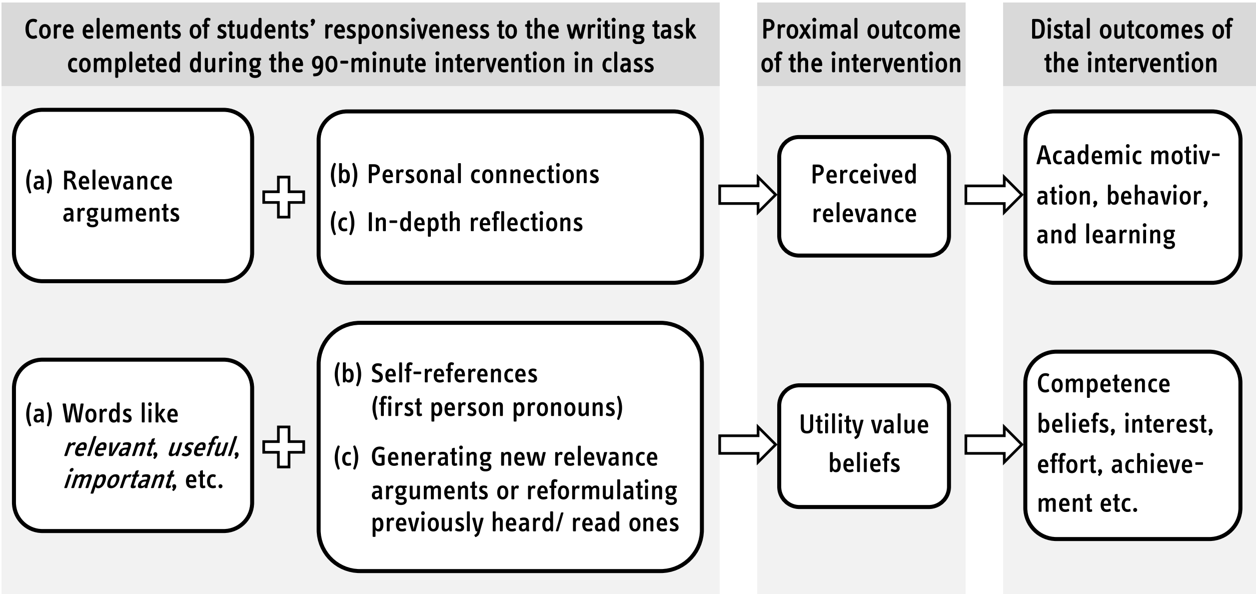



**Figure 1**: Change model (above) and operational model (below) depicting the theorized processes underlying the effectiveness of the MoMa relevance interventions.

**Coding values, coding procedure, and reliability measures.** As indicated in the intervention models, students participating in the MoMa study were supposed to adhere to the three indicators of responsiveness which are presented in Table 1 along with their coding values and reliabilities. All indicators were coded with the values 1 (*low responsiveness*), 2 (*medium responsiveness*), and 3 (*high responsiveness*). The coding values reflect proportions of (a) positive vs. negative relevance arguments for *relevance arguments*, (b) self-references vs. other-references for *personal connections*, and (c) relevance arguments which have been newly generated or reformulated from the intervention material/ transferred to new contexts vs. relevance arguments which have been reproduced from the intervention material for *in-depth reflections* (see Table 1).

Six trained students coded the essays on the indicators of responsiveness using a coding manual (for examples of coded essays see Appendix, Part B). At first, each coder independently coded a randomly chosen set of 10 essays per condition. Intercoder agreement was determined by calculating weighted Cohen's kappa, which is applicable to ratings using ordinal categories and measures the proportion of weighted agreement corrected by chance (Cohen, 1968). Mean weighted Cohen's κ was moderate to almost perfect depending on the coding category and condition: "relevance argumentation" (quotations: $\kappa_{mean}$ = .61 / text: $\kappa_{mean}$ = .92), "personal connections" (.41 / .70), "in-depth argumentation" (.46 / .40) (Landis & Koch, 1977). The coders discussed the inconsistencies and agreed on one common value for each essay and coding category. Subsequently, the rest of the essays (except a random set of 20 essays per condition) was distributed randomly within condition among the coders and coded only once. Four of the coders each coded 244 essays individually and two of the coders each coded 122 essays individually. After half of the individual codings, the randomly chosen 40 essays were coded by all of the coders independently. Intercoder agreements calculated from the second set of multiple-coded essays were substantial to almost perfect for all categories and conditions (see Table 1)—excepting "personal connections" in the quotations condition, for which agreement remained moderate (Landis & Koch, 1977). Finally, the second half of the randomly distributed essays was coded by the coders individually.

Table 1

*Coding values and reliabilities of the indicators of students' responsiveness to the intervention tasks*

| Indicator | Value | | | $\kappa_{mean}$ | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | Quot. | Text |
| Relevance arguments | only negative arguments (i.e., against the utility of math) | ≤ 50% of all arguments are positive (i.e., for the utility of math) | > 50% of all arguments are positive (i.e., for the utility of math) | .66 | .81 |
| Personal connections | only other-references and/or impersonal references, e.g., *they, one, he, his, it, anyone* | < 50% of all personal references are self-references (i.e., first-person pronouns), e.g., *I, me, my* | ≥ 50% of all personal references are self-references (i.e., first-person pronouns), e.g., *I, me, my* | .53 | .81 |
| In-depth reflections | only arguments from the presentation (and / or the quotes[a]) | < 50% of all arguments were new, reformulated or transferred (i.e., not directly copied from the presentation / quotes[a]) | ≥ 50% of all arguments were new, reformulated or transferred (i.e., not directly copied from the presentation / quotes[a]) | .64 | .71 |

*Note.* [a] in the quotations condition only. $\kappa_{mean}$ = weighted Cohen's kappa coefficient (Cohen, 1968). Quot. = quotations condition. Text = text condition.

**Combining the indicators of responsiveness into one index.** Based on theoretical considerations (e.g., Eccles & Wigfield, 2002), a weighting and a summing procedure were used to combine the indicators of responsiveness into a scale ranging from 1 (*lowest responsiveness*) to 11 (*highest responsiveness*) (cf., recommendations by Nelson et al., 2012). As reflecting on the usefulness of mathematics constituted the most basic instruction of the writing task and, theoretically, was considered a prerequisite for the intervention to have positive effects, the indicator "relevance arguments" received a stronger weight than the other two indicators "personal connections" and "in-depth reflections", which were assumed to be equally important (see Figure 1). When students did not stick to the most basic instruction of the writing assignment—as was the case for students writing nonsense or those who wrote about the uselessness of mathe-matics (i.e., low score on relevance arguments)—they received a value of 1 (i.e., the lowest value) on the responsiveness index. Students partly arguing for the utility of mathematics (i.e., medium score on "relevance arguments") were assigned values between 2 and 6 on the index, depending on the sum of the other two indicators. Similarly, students mainly arguing for the utility of mathematics (i.e., high score on "relevance arguments") were assigned values between 7 and 11 on the index (for more information on the index, see Figure 3 in the Results).

**Target psychological process.** As the target psychological process of the relevance intervention, students' math-related utility value beliefs were measured six weeks and five months after the intervention, using a newly developed value instrument (Gaspard, Dicke, Flunger, Schreier et al., 2015). The scale consisted of twelve items (e.g., "I will often need math in my life.", α = .84). All items were measured using a four-point Likert type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*).

**Assessing potential predictors of student responsiveness**

**Stable student characteristics.** Information on students' gender (0 = female, 1 = male) was provided by the teachers. Students' cognitive ability scores were obtained from a figural cognitive ability test (Heller & Perleth, 2000) with 25 items (Cronbach's α = .79). Students' conscientiousness was assessed with a German version of the NEO-FFI (Borkenau & Ostendorf, 1991) in a questionnaire with a 4-point-Likert type scale ranging from 1 (*totally disagree*) to 4 (*totally agree*). The scale consisted of eleven items (e.g., "I am a productive person who always gets the job done.", α = .80).

**Math achievement.** Teachers provided students' results from a curriculum-based standardized math test in the state of Baden-Württemberg taken at the beginning of Grade 9.

**Initial math-related motivation.** Students' math-related competence beliefs were assessed with two scales that were adapted from previous studies (Schwanzer, Trautwein, Lüdtke, & Sydow, 2005; Trautwein & Köller, 2003). Math-related self-concept was measured with five items (e.g., "I am good at math.", α = .93). Homework-related self-efficacy in mathematics was measured with four items (e.g., "When I try hard, I can solve my math homework correctly.", α = .76). Students' math-related intrinsic value beliefs were measured with four items (e.g., "I like doing math.", α = .93; Gaspard, Dicke, Flunger, Schreier, et al., 2015). All items were measured using a four-point Likert type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). Students' initial math-related utility (prior to the intervention) was also explored as predictor of responsiveness, using the newly developed value instrument by Gaspard, Dicke, Flunger, Schreier, et al. (2015).

**Classroom perceptions.** The scale measuring students' perception of classmates' math-related value beliefs consisted of five items (e.g., "Most students in my class consider math as an important subject.", α = .75). Students' perceived disruptions in math class scale was measured with three items (e.g., "Our math lessons are often disrupted.", α = .88). Both scales were taken or adapted from previous studies (e.g., Baumert et al., 2009). All items were answered on a four-point Likert type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*).

### Statistical analyses

**Descriptive statistics.** The means, standard deviations, and *ICCs* at the class level for the responsiveness index, students' individual characteristics and classroom perceptions at T1, and students' utility value beliefs at T2 and T3 are presented per condition in Table 2. The intercorrelations between these variables are accessible in the Appendix, Part C.

Table 2
*Numbers, means, and standard deviations of all variables under investigation*

| | Quotations | | | | Text | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | *M* | *SD* | *ICC* | *N* | *M* | *SD* | *ICC* | *N* | *M* | *SD* | *ICC* |
| Intervention fidelity | | | | | | | | | | | | |
| Responsiveness index | 544 | 8.25 | 2.27 | .06 | 712 | 8.58 | 2.43 | .07 | n/a | n/a | n/a | n/a |
| Predictors (T1) | | | | | | | | | | | | |
| Cognitive ability | 519 | 19.96 | 4.00 | .04 | 681 | 19.99 | 4.22 | .05 | 610 | 19.62 | 4.27 | .02 |
| Conscientiousness | 518 | 2.90 | 0.44 | .05 | 682 | 2.91 | 0.43 | .02 | 608 | 2.91 | 0.44 | .03 |
| Math test score | 517 | 48.67 | 16.49 | .08 | 676 | 49.85 | 18.19 | .21 | 600 | 46.26 | 16.75 | .14 |
| Math self-concept | 515 | 2.76 | 0.79 | .03 | 678 | 2.74 | 0.81 | .04 | 606 | 2.67 | 0.81 | .05 |
| Math HW self-efficacy | 427 | 2.80 | 0.62 | .03 | 599 | 2.72 | 0.62 | .05 | 514 | 2.71 | 0.65 | .04 |
| Math intrinsic value | 515 | 2.31 | 0.84 | .04 | 675 | 2.29 | 0.86 | .09 | 602 | 2.18 | 0.84 | .08 |
| Math utility value | 517 | 2.56 | 0.49 | .05 | 680 | 2.52 | 0.47 | .07 | 607 | 2.52 | 0.49 | .09 |
| Classmates' math valuing | 505 | 1.98 | 0.57 | .24 | 669 | 1.96 | 0.57 | .25 | 602 | 1.85 | 0.52 | .11 |
| Disruptions in math class | 486 | 2.33 | 0.73 | .29 | 652 | 2.35 | 0.77 | .40 | 574 | 2.43 | 0.74 | .30 |
| Outcomes | | | | | | | | | | | | |
| Math utility value (T2) | 530 | 2.64 | 0.50 | .04 | 680 | 2.53 | 0.51 | .08 | 601 | 2.45 | 0.50 | .08 |
| Math utility value (T3) | 516 | 2.60 | 0.49 | .01 | 627 | 2.56 | 0.49 | .11 | 557 | 2.44 | 0.51 | .07 |

*Notes. N* = number; *M* = mean; *SD* = standard deviation; *ICC* = intraclass correlation coefficient relating to the amount of variance explained by differences between classes. HW = homework; T = time; n/a = not applicable.

**Regression analyses.** The association of students' individual characteristics and classroom perceptions with students' intervention responsiveness was analyzed by running regression models in Mplus 7 (Muthén & Muthén, 1998-2012). For each intervention condition, the responsiveness index was regressed on individual student characteristics and classroom perceptions as predictors simultaneously to compare their predictive strength. Standard errors were corrected to account for the nesting of students within classes by using design-based correction of standard errors and test statistics (see McNeish, Stapleton, & Silverman, 2017, for a justification of this approach).

**Model specifications.** Complier-average causal effects (CACE) analyses were run separately for the quotations and the text condition to estimate intervention effects on students' utility value beliefs compared to the control group as a function of students' responsiveness. We estimated multivariate mixture models in Mplus 7 (Muthén & Muthén, 1998-2012), reflecting two classes (compliers vs. noncompliers) where a student who responded to the essay writing activity as intended was considered as a complier (see Appendix, Part D, for a sample syntax). As class membership was only known for the experimental conditions, control group students' responsiveness status was estimated using a) individual student characteristics and classroom perceptions as covariates predicting class membership and b) estimated parameters for the distribution probabilities of compliers and noncompliers in the population (Jo, 2002). Average causal intervention effects were estimated separately for compliers (CACEs) and noncompliers (NCACEs) by regressing students' utility value beliefs at T2 and T3 simultaneously on the intervention dummy variable in the two latent classes (the control group represented the reference group). The nesting of students in classes was controlled by using design-based correction of standard errors and test statistics (see McNeish et al., 2017). The school level was ignored in the analysis, as maximally 4.2 % of the variation in the variables under investigation was explained by differences between schools (after accounting for class-level differences). More information on the CACE models can be found in Nagengast et al. (2018), who also discussed the sensitivity of CACE estimates to different model specifications.

**Cutoff criteria.** Estimating CACEs required dichotomizing the responsiveness index to distinguish between compliers and noncompliers (e.g., Imbens & Rubin, 1997). As recommended by Sagarin et al. (2014), we performed sensitivity analyses using several cutoff values to determine whether the results were robust for different choices of cutoffs. Besides, these sensitivity analyses serve to explore a dose-response relationship between the degree of responsiveness and the strength of intervention effects on students' utility value beliefs.

**Missing data.** In the quotations condition, missing data (see also Table 2) amounted to 3.0 % for the responsiveness index and ranged from 7.5 % to 23.9 % for the predictors (i.e., individual student characteristics and classroom perceptions) as well as from 5.5 % to 8.0 % for the outcomes (i.e., students' utility value beliefs). In the text condition, missing values amounted to 1.1 % for the responsiveness index and ranged from 5.3 % to 16.8 % for the predictors as well as from 5.6 % to 12.9 % for the outcomes. In the control condition, missing data ranged from 3.9 % to 19.1 % for the predictors and from 5.2 % to 12.3 % for the outcomes. Little's (1988) missing completely at random (MCAR) test revealed that missingness was not completely unrelated to the data. Consequently, the full information maximum likelihood (FIML) method which requires the less stringent missing at random (MAR) was used in all analyses (e.g., Graham, 2009). MAR describes a missing data mechanism in which the probability of missingness on a certain variable is related to other observed variables in the dataset but not to the values of the variable itself (Enders, 2010). MAR cannot be tested, but chances of satisfying the MAR assumption can be improved by incorporating correlates of missingness in the analytical model (e.g., Graham, 2009). To make MAR more plausible, individual student characteristics and classroom perceptions were included as covariates in all analyses.

**Effect sizes.** Before running the analyses, all continuous variables were standardized. Consequently, the regression coefficients of the predictor variables can be directly interpreted as measures of the effect size. They indicate the standardized mean difference between students in the control condition and students in the experimental condition (e.g., Tymms, 2004).

## Results

### Students' responsiveness to the writing tasks about relevance

The frequency distributions of the fidelity indicators "relevance arguments", "personal connections", and "in-depth reflections", are presented in Figure 2. Results show that a very small amount of students produced nonsense writings and thus had no values on any of the indicators. Otherwise, in both conditions, a huge majority of the students wrote mostly (or only) arguments about the relevance of mathematics (rather than about its uselessness). Concerning personal connections, most students in the quotations condition used more other-references than self-references. In the text condition, most students used at least the same number of self-references as other-references. As for in-depth reflections, about one third of the students in the quotations condition did not use any new relevance arguments or contexts in their essays. In the text condition, the majority of students used at least one new relevance argument in their essays.

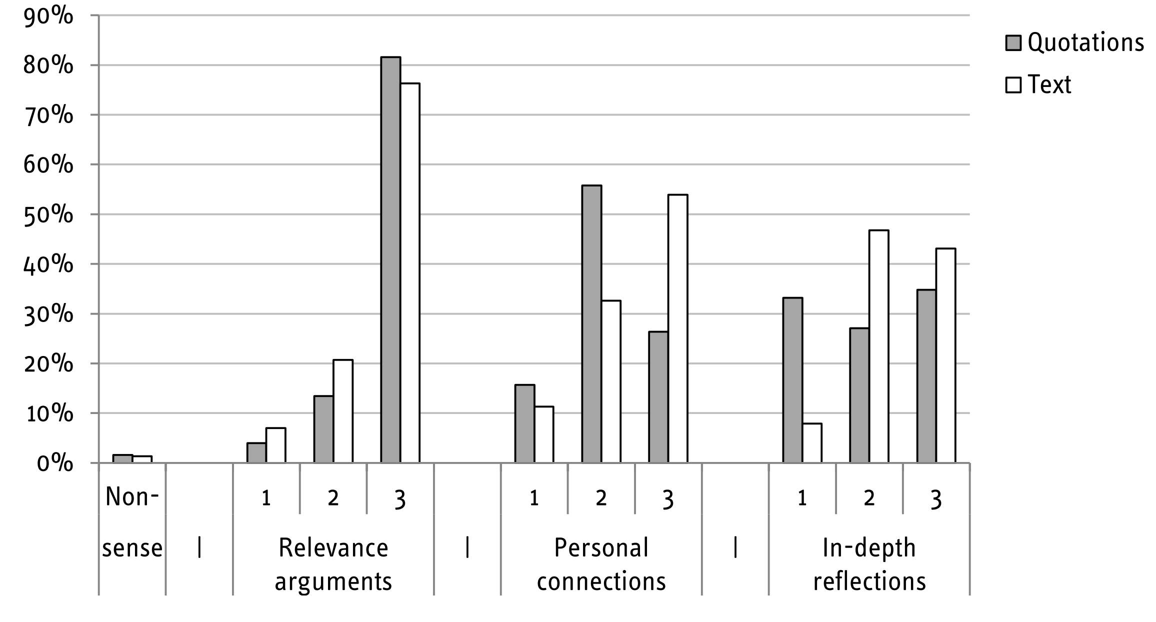
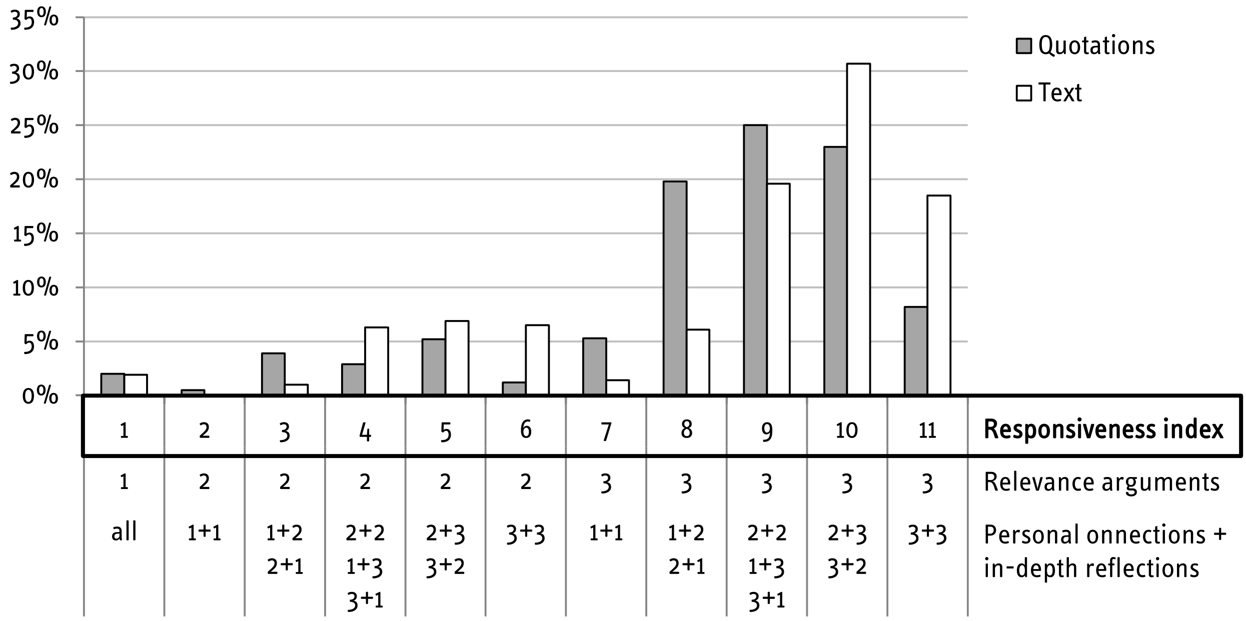



**Figure 2:** Frequency distributions of students' values on the fidelity indicators per condition.

The frequency distributions of the responsiveness index (combining all indicators) are presented in Figure 3. In both the quotations and the text condition, the frequency distributions were skewed to the left indicating overall high levels of responsiveness to the relevance assignments. Only a few students in both conditions received values between 1 and 7 on the total responsiveness index, whereas the majority of students received values between 8 and 11.




Examples:

**Index value of 1:** Students scoring 1 on „relevance"; students producing nonsense

**Index value of 4:** Students scoring 2 on „relevance" and, e.g., 1 on "connections" and 3 on "reflections"

**Index value of 8:** Students scoring 3 on „relevance", and, e.g., 2 on "connections" and 1 on "reflections"

**Index value of 11:** Students scoring 3 on „relevance" and 3 on "connections" and 3 on "reflections"

**Figure 3:** Frequency distributions of students' values on the responsiveness index (and respective scores on the single fidelity indicators of responsiveness) per condition.

**Individual characteristics and classroom perceptions predicting responsiveness**

Results concerning the prediction of students' responsiveness to the writing tasks through students' individual characteristics and classroom perceptions are shown in Table 3. Comparing the relative predictive strength of all predictors, students' conscientiousness was a significant predictor of the responsiveness index in both conditions (quotations: $\beta$ = .09, $p$ = .050; text: $\beta$ = .09, $p$ = .014). In the quotations condition, students' math achievement ($\beta$ = .18, $p$ = .001) and math intrinsic value ($\beta$ = .14, $p$ = .050) predicted the responsiveness index positively, indicating that high-achievers and students who were highly intrinsically motivated for math responded to the quotations assignments significantly better than low-achievers and students with low intrinsic value beliefs of math. In the text condition, students' gender ($\beta$ = -.29, $p$ = .003) emerged as the strongest predictor of students' responsiveness to the relevance essays when controlling for the respective other predictors, indicating that females were more responsive than males. Math motivation also played a role for intervention responsiveness in the text condition: Students with high initial utility value beliefs of math had significantly higher values on the responsiveness index ($\beta$ = .14, $p$ < .001) than students with low initial math utility value. Students' cognitive ability and classroom perceptions were not associated with students' responsiveness in either of the two conditions, controlling for the respective other predictors.

Table 3
*Predicting intervention responsiveness from students' individual characteristics and classroom perceptions*

| | Quotations | | | Text | | |
|---|---|---|---|---|---|---|
| | *β* | *(SE)* | *p* | *β* | *(SE)* | *p* |
| Basic characteristics | | | | | | |
| Gender (1 = male) | -.07 | (.08) | .401 | -.29 | (.10) | .003 |
| Cognitive ability | .02 | (.03) | .593 | .04 | (.04) | .381 |
| Conscientiousness | .09 | (.05) | .050 | .09 | (.04) | .014 |
| Math achievement | | | | | | |
| Test score | .18 | (.06) | .001 | .06 | (.04) | .143 |
| Math motivation | | | | | | |
| Self-concept | -.15 | (.10) | .136 | -.03 | (.05) | .642 |
| HW self-efficacy | .05 | (.06) | .440 | .07 | (.05) | .179 |
| Intrinsic value | .14 | (.07) | .050 | .09 | (.06) | .116 |
| Utility value | .07 | (.05) | .216 | .14 | (.04) | .000 |
| Classroom perceptions | | | | | | |
| Class' math valuing | .06 | (.04) | .206 | .06 | (.04) | .161 |
| Disruptions in math class | -.03 | (.06) | .648 | -.05 | (.04) | .154 |

*Notes.* HW = homework; *β* = standardized regression coefficient; *SE* = standard error; *p* = two-tailed *p*-value.

**Intervention effects on compliers' and noncompliers' math-related utility value**

The intervention effects on students' utility value beliefs in the quotations and the text condition compared to the control group for compliant and noncompliant students 6 weeks (T2) and 5 months (T3) after the intervention are presented in Table 4 and in Figure 4. The regression coefficient of the dummy variable representing the intervention can be directly interpreted as the effect size of the intervention on the outcome as compared to the control condition. As the frequency distributions of students' values on the responsiveness index indicated a sharp increase starting with the value of 8, we first conducted CACE analyses with the cutoff 7/8 classifying students with values of 1 to 7 on the responsiveness index as noncompliers and students with values of 8 to 11 as compliers (Table 4). For our sensitivity analyses (Figure 4), six further CACE models were run using different cutoffs on the responsiveness index to distinguish compliers from noncompliers: three for descending cutoff values (i.e., cutoffs set at the values of 6/7, 5/6, and 4/5) and three for ascending cutoff values (i.e., cutoffs set at the values of 8/9, 9/10, and 10/11). The full models with the results of the sensitivity analyses are accessible in the Appendix, Part E (Tables E1 to E4).

Changes in the cutoff values go along with changes in the size of the groups of compliers and noncompliers, which also affects the statistical power of the analyses (e.g., Stevens, 2012). Therefore, the statistical significance of the intervention effects compared to the control group were tested and are presented in the tables in the Appendix (Part E) but will not be referred to in the results section (cf., Wasserstein & Lazar, 2016). The intervention effects presented in the results section are descriptive and have to be interpreted very cautiously as some of them might not be statistically significant. In contrast, the differences between compliers and noncompliers presented in the results section represent statistically significant effects which were determined using Wald-$\chi^2$-tests (Bakk & Vermunt, 2016).

**CACE analyses using the cutoff 7/8.** As presented in Table 4, based on the cutoff set between the values of 7 and 8 on the responsiveness index, positive effects of the quotations condition on compliers' (C) utility value beliefs compared to the control group emerged at both T2 ($\beta_C$ = .30) and T3 ($\beta_C$ = .23), controlling for students' individual characteristics and classroom perceptions. Noncompliers' (NC) utility value beliefs were also fostered through the quotations condition compared to the control group at both T2 ($\beta_{NC}$ = .29) and T3 ($\beta_{NC}$ = .32). In the text condition, positive effects on compliers' utility value beliefs compared to the control group were observed at both T2 ($\beta_C$ = .19) and T3 ($\beta_C$ = .12). Concerning noncompliers, the text condition had a negative effect on students' utility value beliefs compared to the control group at T2 ($\beta_{NC}$ = -.08) and a positive effect at T3 ($\beta_{NC}$ = .13). According to Wald-$\chi^2$-tests, no statistically significant differences in the strength of the intervention effects at T2 and T3 were observed between compliers and noncompliers at the cutoff between the values of 7 and 8 in either of the two intervention conditions. These results largely correspond with findings by Nagengast et al. (2018), who, however, included students absent at the day of the interventions as noncompliers in their analyses.

**Sensitivity CACE analyses.** As shown in Figure 4, the sensitivity analyses revealed positive effects of the quotations condition on both compliers' and noncompliers' utility value beliefs at all other cutoffs for both T2 ($\beta_C$ = .28 to 1.33; $\beta_{NC}$ = .08 to .35) and T3 ($\beta_C$ = .20 to 1.33; $\beta_{NC}$ = .17 to .45), controlling for students' individual characteristics and classroom perceptions. Interestingly, Wald-$\chi^2$-tests showed that when the cutoff was set above the values of 7 and 8, the intervention effects were significantly stronger for compliers than for noncompliers at T2, indicating that the most responsive students benefitted most from the quotations condition at T2.

The sensitivity analyses for the text condition also showed positive intervention effects on compliers' utility value beliefs at all other cutoffs at both T2 ($\beta_C$ = .18 to .22) and T3 ($\beta_C$ = .07 to .20). However, for noncompliers, the direction of the effects depended on the cutoff values: When the cutoff was set at values below 7 and 8, negative intervention effects emerged at T2 ($\beta_{NC}$ = -.47 to -.13) and, partially, at T3 ($\beta_{NC}$ = -.22 to .10), whereas for cutoffs above the values of 7 and 8, positive intervention effects were obtained at both T2 ($\beta_{NC}$ = .06 to .10) and T3 ($\beta_{NC}$ = .14 to .17). Wald-$\chi^2$-tests revealed that the differences in the strength of the intervention effects between compliers and noncompliers at the lower cutoff values were at least marginally statistically significant at T2, indicating that only students with a minimal degree of responsiveness benefitted from the text condition.

Table 4

*Intervention effects on students' math utility value depending on students' intervention responsiveness: Cutoff on responsiveness index between values of 7 (NC max. value) and 8 (C min. value)*

| | Quotations | | | | | | Text | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 188 (16 %) | | | | | | 241 (18 %) | | | | | |
| Complier | 1008 (84 %) | | | | | | 1114 (82 %) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | *(SE)* | *p* | $\beta_{NC}$ | *(SE)* | *p* | $\beta_{NC}$ | *(SE)* | *p* | $\beta_{NC}$ | *(SE)* | *p* |
| Intervention | .29 | (.20) | .142 | .32 | (.21) | .124 | -.08 | (.20) | .673 | .13 | (.32) | .681 |
| Covariates | | | | | | | | | | | | |
| Gender (1 = male) | .19 | (.18) | .296 | .11 | (.16) | .480 | -.12 | (.12) | .346 | .05 | (.14) | .721 |
| Cognitive ability | -.20 | (.07) | .002 | -.09 | (.08) | .263 | -.09 | (.06) | .126 | -.14 | (.08) | .088 |
| Conscientiousness | .07 | (.07) | .321 | .06 | (.07) | .353 | -.05 | (.08) | .531 | -.10 | (.08) | .211 |
| Test score | .11 | (.06) | .075 | .04 | (.09) | .666 | -.01 | (.07) | .929 | .08 | (.09) | .362 |
| Self-concept | .07 | (.10) | .522 | .15 | (.12) | .198 | .05 | (.12) | .702 | .08 | (.17) | .658 |
| HW self-efficacy | -.07 | (.11) | .519 | -.15 | (.09) | .085 | .08 | (.07) | .276 | .09 | (.13) | .513 |
| Intrinsic value | .05 | (.08) | .562 | -.08 | (.14) | .546 | .01 | (.11) | .922 | -.11 | (.11) | .309 |
| Utility value | .44 | (.09) | .000 | .50 | (.10) | .000 | .66 | (.11) | .000 | .53 | (.11) | .000 |
| Class' math valuing | .02 | (.07) | .795 | .03 | (.10) | .790 | -.01 | (.07) | .938 | .17 | (.10) | .066 |
| Disruptions in class | .02 | (.06) | .751 | .06 | (.09) | .520 | .01 | (.07) | .914 | -.05 | (.07) | .501 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | .30 | (.07) | .000 | .23 | (.07) | .001 | .19 | (.07) | .010 | .12 | (.11) | .266 |
| Covariates | | | | | | | | | | | | |
| Gender (1 = male) | .08 | (.05) | .151 | .06 | (.06) | .251 | .04 | (.08) | .593 | .09 | (.07) | .173 |
| Cognitive ability | .02 | (.02) | .493 | .03 | (.03) | .404 | -.02 | (.03) | .477 | .00 | (.03) | .925 |
| Conscientiousness | -.04 | (.03) | .249 | -.06 | (.03) | .054 | -.03 | (.03) | .259 | -.03 | (.03) | .448 |
| Test score | -.03 | (.03) | .336 | -.04 | (.04) | .294 | .06 | (.03) | .038 | .04 | (.03) | .265 |
| Self-concept | -.07 | (.06) | .227 | .05 | (.05) | .290 | -.01 | (.06) | .848 | -.02 | (.06) | .792 |
| HW self-efficacy | .04 | (.04) | .417 | -.03 | (.05) | .544 | .00 | (.04) | .964 | .00 | (.05) | .944 |
| Intrinsic value | .06 | (.04) | .132 | .06 | (.04) | .169 | .03 | (.05) | .602 | .05 | (.04) | .281 |
| Utility value | .70 | (.04) | .000 | .61 | (.04) | .000 | .64 | (.05) | .000 | .57 | (.05) | .000 |
| Class' math valuing | .03 | (.04) | .410 | .05 | (.03) | .134 | .05 | (.04) | .302 | .07 | (.04) | .088 |
| Disruptions in class | -.02 | (.03) | .605 | .01 | (.04) | .773 | -.04 | (.03) | .200 | -.02 | (.03) | .458 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | 0.00 | (1) | .990 | 0.14 | (1) | .706 | 1.28 | (1) | .258 | 0.00 | (1) | .986 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

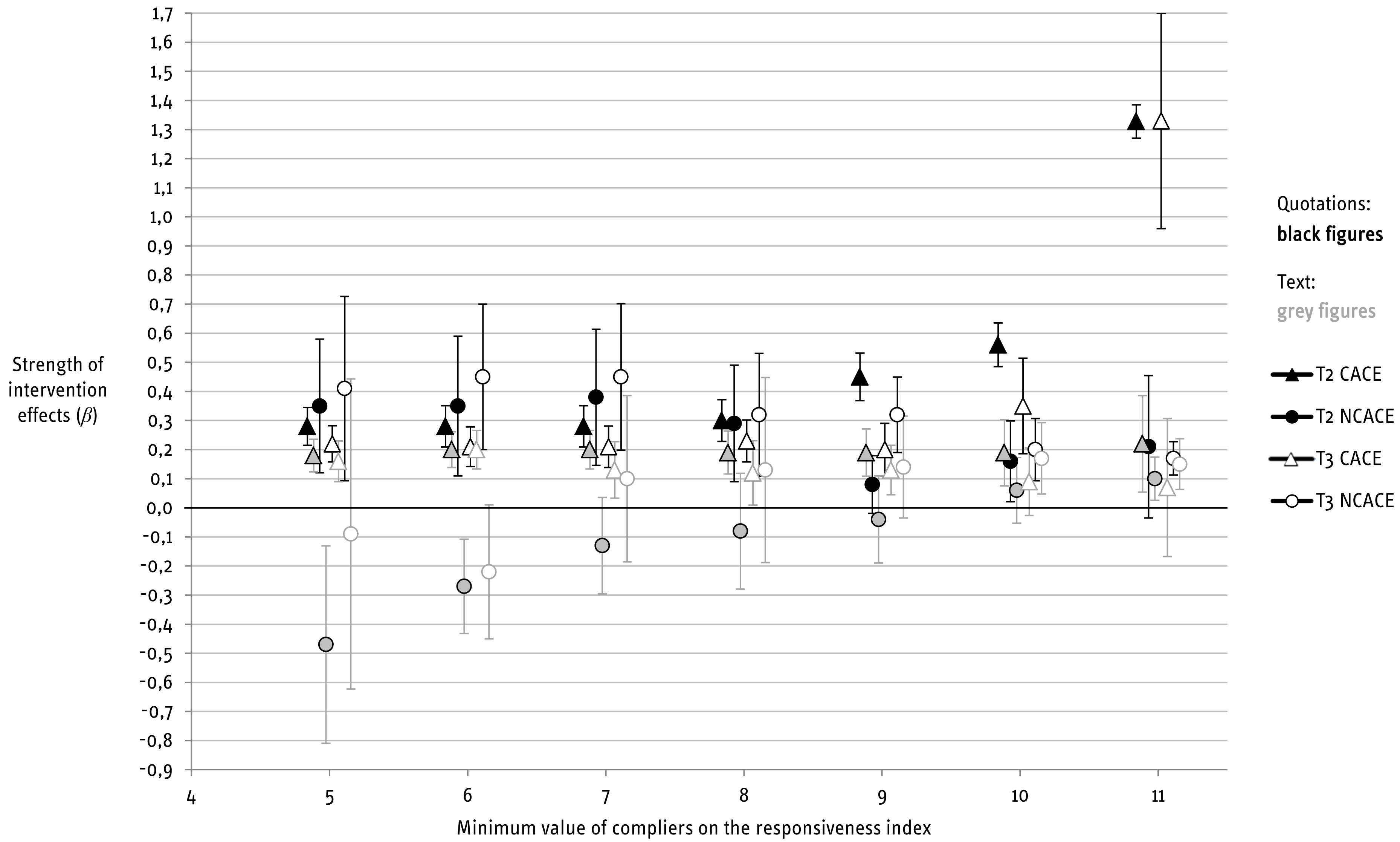



**Figure 4:** Intervention effects on students' math utility value depending on different degrees (cutoff values) of students' responsiveness per condition.

## Discussion

Who sticks to the instructions of the writing tasks in educational interventions? And does it matter after all? Although writing tasks are a common tool used in psychological interventions to change students' personal beliefs (Yeager & Walton, 2011), comprehensive studies assessing if intervention processes are related to psychological processes in ways that support the theory of change are missing. In this study, we sought to fill that gap in the literature by investigating whether students did what they were asked to do during the intervention (i.e., responsiveness), the characteristics of responsive students, and how the degree of responsiveness was related to a change in students' utility value beliefs—the target psychological process of relevance interventions. Based on a sample of 1961 ninth-graders participating in the MoMa project and 1280 essays written during the MoMa interventions, the current study found girls, highly conscientious students, high achievers, and students with high math-related motivation to be most responsive to written intervention tasks about the relevance of mathematics. When students evaluated quotations about the relevance of math, math utility value of both highly and lowly responsive students was raised through the intervention compared to the control group. When students wrote a text about the personal relevance of math, highly responsive students benefitted from the intervention, whereas lowly responsive students' math utility value was not raised. To the contrary, negative intervention effects on students' utility value beliefs were found for the students with the lowest levels of responsiveness. The degree to which intervention responsiveness mattered thus depended on the relevance task.

### The rationale of relevance interventions: different tasks, different intervention processes

Prior analyses of the MoMa dataset have shown that evaluating quotations about the relevance of mathematics led to stronger effects on students' math-related motivation, effort, and achievement than writing a text about the personal relevance of mathematics, and that girls benefited more than boys from the text condition (Brisson et al., 2017; Gaspard, Dicke, Flunger, Brisson, et al., 2015). The current study introduced and tested a systematically derived theoretical framework for assessing and analyzing responsiveness, so as to learn more about the intervention processes leading to changes in students' utility value beliefs and potentially contributing to overall differences in the effectiveness of the two conditions. In contrast to Nagengast et al. (2018) who compared the results of different approaches to estimate the effectiveness of the MoMa interventions with and without including responsiveness (ITT analyses vs. CACE models; e.g., Boruch, 1997; Sagarin et al., 2014), we developed a theory-driven framework to assess student responsiveness and tested it by examining intervention processes depending on different degrees of responsiveness. These sensitivity analyses provided important insights into what happened with the most and the least responsive students after the intervention, and how these effects differed between the two intervention conditions. Through excluding students absent during the day of the intervention from the analyses (instead of treating them as noncompliers as did Nagengast et al., 2018) the current results for the least responsive students specifically refer to those who did complete the relevance task but not as expected. These students deserve particular attention to make sure that the interventions did not produce any unexpected negative effects.

Based on the assumed change model, students' responsiveness to the MoMa intervention assignments was assessed by coding the degree of positive argumentation, personal connections, and in-depth reflections about relevance in students' essays (e.g., Eccles et al., 1983). For theoretical reasons, the codings were combined into an overall responsiveness index by giving a stronger weight to the degree of positive argumentation (which was descriptively similar in both conditions) than to the other two fidelity indicators (cf. recommendations by Nelson et al., 2012). As a result, overall responsiveness was similarly high in both conditions (although students in the text condition had descriptively higher values on personal connections and in-depth reflections than students in the quotations condition). Therefore, student responsiveness per se cannot explain the differences in the strength of the two intervention approaches. Indeed, a closer look at the predictors of students' responsiveness and at the effects on students' math-related utility value as a function of students' degree of responsiveness rather indicates that different intervention processes are at work in these two conditions.

**Writing a text about relevance: new insights through analyzing responsiveness.** In the text condition, students with high initial utility value beliefs, girls, and highly conscientious students had the highest levels of responsiveness to the relevance task, holding other individual and classroom characteristics constant. Controlling for these individual differences, stronger intervention effects on students' math-related utility value compared to the control group were observed for more responsive students than for less responsive students, whose utility value beliefs could not be fostered. In contrast, not responding to the responsiveness criteria might even have caused negative effects on students' utility value beliefs. Despite a minority of less than 20 %, the noncompliers probably contributed to the general pattern of results showing the text condition to be less effective overall than the quotations condition. Which insights do these findings provide into the processes that make the text condition trigger a change in students' motivational beliefs—or not?

First, the intervention effects might pertain to positive and negative self-reinforcing processes (Yeager & Walton, 2011). More precisely, students with low initial math utility value seemed to have (partially) argued against the relevance of mathematics in their writing assignments; maybe they also realized how hard it is for them to come up with several utility arguments. As a consequence, the intervention might have caused them to think of mathematics to be even more useless than before. In contrast, students who initially had high utility value beliefs possibly had more concrete ideas about the usefulness of mathematics, and writing them down might have reinforced their positive beliefs about mathematics.

Second, boys might have disliked the text writing task more than girls and thus responded less well to the task. Indeed, writing an essay resembles typical activities done in language subjects—in which boys have also been found to comply less with homework than girls

(Trautwein & Lüdtke, 2009; Trautwein et al., 2006). The current results thus indicate that girls' high degrees of intervention responsiveness might have contributed to the gender effects found in the text condition (for more detailed analyses, see Nagengast et al., 2018).

Finally, the text condition promoted students' utility value beliefs to a lesser degree than the quotations condition—even for the most responsive students. The current findings might also be interpreted in a way that the effectiveness of the MoMa interventions might have resulted from an interaction of students' responsiveness to the writing task and their reaction to the initial presentation of the utility of mathematics. In the text condition, the positive effect of the presentation might have been undermined by the rather difficult subsequent task of having to write the essays. Only students who responded well to the writing assignment might have benefitted from the positive effect of the presentation. In other words, it might be necessary for the writing task to be easy enough for all students so that the input on relevance given in the presentation can unfold its full potential on the students.

**Evaluating quotations about relevance: extending the intervention models.** Conscientious students, high achievers, and students with high math-related intrinsic value beliefs responded best to the quotations-based writing assignment, controlling for all other individual and classroom characteristics. Students' responsiveness was also of importance in the quotations condition, as the strongest intervention effects were observed for the most responsive students. However, the degree of responsiveness did not matter for the direction of the intervention effects: Holding students' individual differences constant, math-related utility value of both responsive and nonresponsive students was fostered through the quotations task. On the one hand, these results underline the strength of this intervention approach to induce positive effects regardless of students' degree of responsiveness. On the other hand, the index—as created from the degree of positive argumentation, personal connections, and in-depth reflections—seems to say very little about the processes leading to a change in students' motivational beliefs through the quotations task. Which elements might be missing in the change model and/or the operational model?

First, students who read the quotations were provided more relevance information than students in the text condition. The rather simple fidelity criterion "reading the quotations" might thus already have contributed to the strength of the quotations approach.

Second, following assumptions made in identity-based motivation (Oyserman & Destin, 2010), the intervention might have had a particularly strong effect when students personally connected to the interviewees and their situations. In this case, the stronger use of other-references than self-references (used to measure the fidelity indicator "personal connections") would probably not constitute an indicator of a low degree of responsiveness, because students strongly connecting with the interviewees' quotations might more frequently have used third-person pronouns in their essays (Pennebaker et al., 2003). Measuring the degree of identification and emotional closeness with others however requires assessing a variety of words and expressions—a complex task which is typically realized with computerized methods that cannot easily be applied to the handwritten essays (cf., Harackiewicz et al., 2016; Tausczik & Pennebaker, 2010).

Third, the task characteristics might have triggered cognitive engagement and in-depth reflections in a way that could only marginally be assessed through coding the use of self-generated vs. reproduced (i.e., previously heard or read) relevance arguments. Indeed, reading and evaluating the personal importance of relevance quotations required students to make judgments based on their own standards and those defined by the task—deep-level cognitive processes, which are less easily triggered in mere production tasks such as the text writing condition (Krathwohl, 2002), and which are difficult to capture reliably in students' responses to paper and pencil tasks as used in the MoMa interventions.

**Paving the way for relevance interventions to enter educational practice**

Throughout secondary school, students' motivational beliefs are declining—in particular their utility value beliefs (e.g., Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2017; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). Prior research has shown that relevance interventions—even if they are as short as 90 minutes in class—are a powerful tool to halt this decline in students' motivation and thereby support students' academic interest, feelings of competence, effort (as observed by teachers), grades, and test-based achievement in real-life classroom settings (for an overview, see e.g., Rosenzweig & Wigfield, 2016). The current investigation extends prior relevance intervention research by showing that high intervention fidelity on the side of the participants is desirable: In both conditions, the strongest intervention effects on students' math-related utility value were found for the students with the highest degrees of responsiveness to the intervention activity. In the quotations condition intervention effects were as large as $d$ = 1.33 for the most responsive students—which is by far larger than the expectable average impact of motivation interventions implemented at the high school level ($d$ =.42; Lazowski & Hulleman, 2016). Though small applied to conventional standards (Cohen, 1969), even the effect size of $d$ = .22 for the most responsive students in the text condition can be considered a meaningful improvement in students' motivation in practical terms given the brevity of the intervention and the small sample size (see also Gaspard, Dicke, Flunger, Schreier et al., 2015).

The current results not only indicate that the effectiveness of short classroom-based relevance interventions might be further boosted by improving the implementation of intervention activities for students (Nelson et al., 2012). They also provided insight into the characteristics of students who are most likely to be nonresponsive to writing activities used in relevance interventions, for example, lowly conscientious students or male students. Future researchers should use these findings to investigate ways to enhance students' intervention responsiveness, to optimize the designs of relevance interventions, and thereby pave their way for entering educational practice. Interviewing students who are lower in responsiveness might be helpful in exploring reasons of students' negative reactions to intervention activities. If the assignment was not intelligible enough for them, providing more scaffolding might be necessary, for example, by tying oral or written instructions closer to the criteria of responsiveness (e.g., by explicitly asking students to use first-person pronouns and to come up with own relevance arguments). If the intervention activity was not attractive to students, changes to the instructions that reduce reactance need to be developed and tested.

As in the quotations-based relevance intervention, even a minimal degree of responsiveness has been shown to lead to sustained and meaningful effects on students' utility value beliefs, this approach—preceded by a psychoeducational presentation—seems suitable for implementation in various student groups. Indeed, researchers and educators promote the use of evidence-based relevance-enhancing teaching approaches, particularly in science-related subjects (e.g., Davis & McPartland, 2012; Osborne & Dillon, 2008). To prepare for scaling-up in educational practice, it is advisable to test the quotations approach within diverse student samples, to include teachers in the implementation process (Cohen & Loewenberg Ball, 2007), and to (re)investigate the fidelity criteria used in this research as well as new responsiveness criteria.

In contrast, when implementing a free text writing assignment about relevance, researchers and educators should be aware that low degrees of responsiveness might have detrimental effects on students' psychological processes. Although the negative effect of $d$ = -.47 on utility value for the least responsive students was statistically nonsignificant ($p$ = .165) and applied to only 6 % of the students, these observed tendencies are meaningful from a practical perspective (cf., Wasserstein & Lazar, 2016): Any potential risks going along with implementing an intervention in educational practice should be avoided—no matter how small the amount of affected students. Certain students were found to be more at risk of responding less well to the intervention task than others, for example, boys. Because girls in Western countries are underrepresented in careers related to science, technology, engineering, and mathematics (e.g., OECD, 2012), one could argue that they are more "in need" than boys. Girls responded well to the text assignment, which indicates the potential of this approach to specifically support girls in finding relevance in mathematics. Nevertheless, it remains problematic that students with low initial math-related utility value had difficulty to fulfil the responsiveness criteria of the free text writing task. It would therefore be desirable to test if this approach can be adapted in a way to support all students, including boys and girls and students with low initial utility value beliefs.

**Advancing research on psychological interventions in education**

Many field-based psychological interventions in education, even if brief in nature, can be effective in raising important academic outcomes (Lazowski & Hulleman, 2016). Hence, the scientific interest in publishing educational interventions remains high (see Graham, 2015). At the same time, the demand for expertise to successfully adapt such interventions to diverse educational contexts is growing—otherwise, it is difficult to replicate their effects (e.g., Yeager & Walton, 2011). Indeed, in educational settings, the source of the effectiveness or noneffectiveness of interventions may be blurred by the number of factors which cannot be kept constant across classrooms (Weiss, Bloom, & Brock, 2014). Using theoretically sound research designs is therefore crucial to deal with unobservable variations across classrooms (e.g., Rubin, 1974). Variation across classrooms in program implementation, in contrast, can be made at least partially observable through measures of fidelity.

The current study showed that assessing and analyzing indicators of intervention fidelity is essential: Investigating the antecedents and effects of intervention fidelity can help to provide an empirical account on the role of concrete core elements concerning direct effects (and potential differential effects) of field-based interventions. Such an empirical understanding of the intervention processes underlying a change in students' beliefs and thereby contributing to the effectiveness of educational interventions is essential to advance psychological theorizing and to enhance the "psychological precision" (Walton, 2014, p. 74) of classroom interventions. Intervention fidelity needs to be researched within diverse learning contexts in order to enable an evidence-based adaptation of specific intervention components to specific educational settings. In addition, studies on intervention fidelity can help to inform researchers and educators about any potential risks associated with inappropriate implementation of, or particpants' nonresponsiveness to, a certain intervention program. Fidelity studies should thus not be an exception, but the rule to go along with any experimental research in the field. Starting from the MoMa interventions as an example, the current study serves as a blueprint of how to do high quality fidelity assessment within an education field experiment.

**Limitations and suggestions for future research**

As to all research studies, several limitations apply to the current investigation pertaining to, for example, the specific sample investigated in the current study. To ensure generalizability, the current results need to be replicated with other samples including students in other education systems as well as German students in non-academic track schools (e.g., vocational track schools). Depending on the focus of the education system or school track, the contents of the intervention material would probably have to be adapted. In addition, nonresponsiveness to relevance interventions might constitute a larger problem at schools which—in contrast to German academic track schools—feature high rates of students at risk (e.g., students with severe cognitive or behavioral problems).

A second critical aspect concerns the fidelity measures. The responsiveness index constitutes a strength of the current study: Several fidelity indicators were assessed and combined based on theoretical considerations, creating an elaborate measure of students' responsiveness to the writing assignments (cf., Nelson et al., 2012). Because of their brevity, it was however impossible to use the same theory-driven coding criteria for the writing activities filled out at home (intervention boosters, see Appendix, Part F). Therefore, the current study was focused on students' responsiveness to the in-class writing tasks.

Third, as recommended by Nelson et al. (2012), we focused on the importance of overall responsiveness to the relevance writing tasks by using one index. To disentangle the importance of each indicator for the intervention effects, it would be necessary to conduct further investigations. For example, by experimentally manipulating the single indicators (as Hulleman et al., 2017, did for "personal connections"), the contribution of each indicator to the effectiveness of relevance interventions could be analyzed in more detail.

Fourth, to enable a comparison between the two relevance intervention conditions, the same intervention models and, therefore, the same indicators were used to assess students' responsiveness to the different writing tasks. However, the coding categories did not seem to be equally straightforward in both conditions, as reliability measures were lower in the quotations condition than in the text condition. In particular, the coding of the pronouns (first-person vs. third-person) in students' essays reached almost perfect intercoder agreement in the text condition but moderate agreement in the quotations condition. Consequently, the results referring to the effects of students' responsiveness can be considered as less robust for the quotations condition than for the text condition. Moreover, through a consistent choice of fidelity indicators, it was impossible to go more into depth with regards to further criteria which might have been specific to the effectiveness of only one of the two conditions. Future studies could investigate the importance of other fidelity indicators such as students' identification with the interviewees (quotations condition) or cognitive engagement in relevance interventions, for example, by combining different media (e.g., reading vs. hearing quotations) with computerized tasks or think-aloud methods (Ericsson & Simon, 1980).

Fifth, the complier-average causal effects models suggested by Jo (2002) were a valuable tool for the purpose of the current investigation. The drawback of CACE analyses is that they typically require the use of a binary variable to determine students' responsiveness status (i.e., compliers vs. noncompliers, Sagarin et al., 2014), resulting in a loss of valuable information on the degree of students' responsiveness when analyzing causal intervention effects. Instead of modelling responsiveness as a continuous variable, the sensitivity analyses using different cutoff values were conducted to explore the dose-response relationship between the degree of responsiveness and the intervention effects. It would be an important task for methodological researchers to develop alternative analytical approaches taking different degrees of responsiveness into account when estimating causal intervention effects.

Finally, the writing assignments constituted the core element of the MoMa relevance interventions and thus were in the focus of the current investigation. Nevertheless, students' experiences during the introductory psychoeducational presentation (e.g., their cognitive engagement while learning about examples for the utility of math) might have also affected the intervention effects. Students' experiences during such a pre-writing part of the intervention could be taken into account in future research, for example, by using experience sampling methods (Rosenzweig & Wigfield, 2016) or observational methods (cf., Fredricks et al., 2004).

## Conclusions

Relevance interventions—which are extremely cost effective (Yeager & Walton, 2011) and easy to implement across different academic settings and domains (Lazowski & Hulleman, 2016)—show a huge potential to raise important learning outcomes (Durik et al., 2015). The results of the present study on student responsiveness imply that when designing written relevance intervention tasks aimed for implementation in real-life educational settings, is it important to consider that individual student characteristics such as conscientiousness, gender, and domain-specific motivation may determine how well the students follow the instructions of the assignments. Depending on the type of intervention task, students' responsiveness may in turn affect the effectiveness of the intervention to a greater or lesser extent. In both approaches of the MoMa relevance interventions, the most responsive students reported the largest increases in motivational beliefs compared to the controls. However, only in the text condition—an approach used in numerous other relevance intervention programs (e.g., Hulleman & Harackiewicz, 2009)—students' responsiveness mattered for the direction of the intervention effects and actually undermined the least responsive students' motivation. Using this knowledge to further improve the theories and designs of relevance interventions might help to eventually pave the way for relevance interventions to enter educational practice at a larger scale.

## References

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 20–31. https://doi.org/10.1080/10705511.2014.955104

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., & Krauss, S. (2009). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung Nr. 83 [Documentation of the data collection instruments of the COACTIV study. Materials of Research in Education Nr. 83]*. Berlin: Max-Planck-Institut für Bildungsforschung.

Borkenau, P., & Ostendorf, F. (1991). Ein Fragebogen zur Erfassung fünf robuster Persönlichkeitsfaktoren [A questionnaire for assessing five robust personality factors]. *Diagnostica*, *37*(1), 29–41.

Boruch, R. F. (1997). *Randomized experiments for planning and evaluation*. Thousand Oaks, CA: Sage.

Brisson, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Nagengast, B., & Trautwein, U. (2017). Short intervention, sustained effects: Promoting students' math competence beliefs, effort, and achievement. *American Educational Research Journal*, *54*(6), 1048–1078. https://doi.org/10.3102/0002831217716084

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility value information. *Motivation Science*, *1*(1), 47–71. https://doi.org/10.1037/mot0000015

Cohen, D. K., & Loewenberg Ball, D. (2007). Educational innovation and the problem of scale. In B. Schneider & S.-K. McDonald (Eds.), *Scale-up in education: Ideas in principle* (pp. 19–36). Plymouth, UK: Rowman & Littlefield Publishers.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, *70*(4), 213–220. https://doi.org/10.1037/h0026256

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. San Diego, CA: Academic Press.

Cooper, H. (1989). *Homework*. New York: Longman.

Cunningham, D., & Duffy, T. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 170–198). New York: Macmillan Publishers.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*(1), 23–45. https://doi.org/10.1016/S0272-7358(97)00043-3

Davis, M. H., & McPartland, J. M. (2012). High school reform and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 515–539). New York: Springer.

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

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3-4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). New York: Freeman.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251. https://doi.org/10.1037/0033-295x.87.3.215

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97–132). New York: Springer.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, *17*(5), 478–493. https://doi.org/10.1016/j.learninstruc.2007.09.001

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, *51*(9), 1226–1240. https://doi.org/10.1037/dev0000028

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, *107*(3), 663–677. https://doi.org/10.1037/edu0000003

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, *48*, 67–84. https://doi.org/10.1016/j.cedpsych.2016.09.003

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Graham, S. (2015). Inaugural editorial for the Journal of Educational Psychology. *Journal of Educational Psychology*, *107*(1), 1–2. https://doi.org/10.1037/edu0000007

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology*, *111*(5), 745–765. https://doi.org/10.1037/pspp0000075

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klasse [Cognitive Ability Test for Grades 4 through 12]*. Göttingen, Germany: Hogrefe.

Hohmann, A. A., & Shear, M. K. (2002). Community-based intervention research: Coping with the "noise" of real life in study design. *American Journal of Psychiatry*, *159*(2), 201–207. https://doi.org/10.1176/appi.ajp.159.2.201

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88–110. https://doi.org/10.1080/19345740802539325

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*(2), 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*(4), 880–895. https://doi.org/10.1037/A0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology*, *109*(3), 387–404. https://doi.org/10.1037/edu0000146

Husman, J., Nelson, K., & Cheng, K. (2017). *Intervening to promote engineering student success in technical, nonmajor, required barrier courses*. Poster presented at the 17th Biennial EARLI Conference. Tampere, Finnland.

Imbens, G. W., & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics*, *25*(1), 305–327.

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*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jo, B. (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics*, *27*(4), 385–409. https://doi.org/10.3102/10769986027004385

Juvonen, J., Espinoza, G., & Knifsend, C. (2012). The role of peer relationships in student academic and extracurricular engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 387–492). New York: Springer.

Karabenick, S. A., Albrecht, J., & Rausch, N. (2017). *Examining the self-generation of subjective task value: A mixed-methods investigation*. Paper presented at the 17th Biennial EARLI Conference. Tampere, Finnland.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Landis, J. R., & Koch, G. G. (1977). Measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954–969. https://doi.org/10.1037/0003-066X.41.9.954

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice*, *41*(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, *59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. https://doi.org/10.1037/met0000078

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, *24*(3), 315–340. https://doi.org/10.1177/109821400302400303

Murrah, W. M., Kosovich, J. J., & Hulleman, C. S. (2017). A framework for incorporating intervention fidelity in educational evaluation studies. In G. Roberts, S. Vaughn, T. Beretvas,

& V. Wong (Eds.), *Treatment fidelity in studies of educational intervention* (pp. 39–60). New York: Routledge.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide, seventh edition*. Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Brisson, B. M., Hulleman, C. S., Gaspard, H., Häfner, I., & Trautwein, U. (2018). Learning more from educational intervention studies: Estimating complier average causal effects in a relevance intervention. *The Journal of Experimental Education*, *86*(1), 105–123. https://doi.org/10.1080/00220973.2017.1289359

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The Journal of Behavioral Health Services & Research*, *39*(4), 374–396. https://doi.org/10.1007/s11414-012-9295-x

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

OECD. (2012). *Closing the gender gap: Act now*. Paris. Retrieved from http://dx.doi.org/10.1787/9789264179370-en

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*. London: The Nuffield Foundation.

Oyserman, D., & Destin, M. (2010). Identity-based motivation: Implications for intervention. *The Counseling Psychologist*, *38*(7), 1001–1043. https://doi.org/10.1177/0011000010374775

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*(1), 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Reschly, A. L., & Christenson, Sandra, L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–20). New York: Springer.

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *Elementary School Journal*, *100*(5), 443–471. https://doi.org/10.1086/499650

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM Motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist*, *51*(2), 146–163. https://doi.org/10.1080/00461520.2016.1154792

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/H0037350

Sagarin, B. J., West, S. G., Ratnikov, A., Homan, W. K., Ritchie, T. D., & Hansen, E. J. (2014). Treatment noncompliance in randomized experiments: Statistical approaches and design issues. *Psychological Methods*, *19*(3), 317–333. https://doi.org/10.1037/met0000013

Schreier, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Lüdtke, O.,. . . Trautwein, U. (2014). Der Wert der Mathematik im Klassenzimmer: Die Bedeutung relevanzbezogener Unterrichtsmerkmale für die Wertüberzeugungen der Schülerinnen und Schüler [The value of mathematics in the classroom: The importance of a relevance-oriented learning environment for students' value beliefs]. *Zeitschrift für Erziehungswissenschaft / German Journal of Educational Studies*, *17*(2), 225–255. https://doi.org/10.1007/s11618-014-0537-y

Schwanzer, A. D., Trautwein, U., Lüdtke, O., & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener [Development of an instrument to assess young adults' self-concept]. *Diagnostica*, *51*(4), 183–194. https://doi.org/10.1026/0012-1924.51.4.183

Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering against identity threat. *Personality and Social Psychology Bulletin*, *39*(5), 663–676. https://doi.org/10.1177/0146167213480816

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th). London: Routledge.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Trautwein, U., & Köller, O. (2003). Was lange währt, wird nicht immer gut: Zur Rolle selbstregulativer Strategien bei der Hausaufgabenerledigung [Time investment does not always pay off: The role of self-regulatory strategies in homework execution]. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology*, *17*(3), 199–209. https://doi.org/10.1024//1010-0652.17.34.199

Trautwein, U., & Lüdtke, O. (2007). Students' self-reported effort and time on homework in six school subjects: Between-students differences and within-student variation. *Journal of Educational Psychology*, *99*(2), 432–444. https://doi.org/10.1037/0022-0663.99.2.432

Trautwein, U., & Lüdtke, O. (2009). Predicting homework motivation and homework effort in six school subjects: The role of person and family characteristics, classroom factors, and school track. *Learning and Instruction*, *19*(3), 243–258. https://doi.org/10.1016/j.learninstruc.2008.05.001

Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology*, *97*(6), 1115–1128. https://doi.org/10.1037/a0017048

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology*, *98*(2), 438–456. https://doi.org/10.1037/0022-0663.98.2.438

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55–66). London: National Foundation for Educational Research.

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*(1), 73–82. https://doi.org/10.1177/0963721413512856

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on *p*-values: Context, process, and purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Weiss, M. J., Bloom, H. S., & Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects. *Journal of Policy Analysis and Management*, *33*(3), 778–808. https://doi.org/10.1002/pam.21760

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1–35. https://doi.org/10.1016/J.Dr.2009.12.001

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg (Ed.), *Handbook of child psychology: Volume 3: Social, emotional, and personality development* (6th ed., pp. 933–1002). Hoboken, NJ: John Wiley & Sons.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301. https://doi.org/10.3102/0034654311405999

# General discussion

## General discussion

Mathematical skills have long been considered useful in everyday life (Klein & Schimmack, 1907). In the 21st century, mathematical skills are not only useful but needed for active participation in society and in numerous professions (e.g., Renn et al., 2012). Results of recent research indicate that secondary school students have trouble seeing the usefulness of mathematical skills in everyday life (e.g., Reiss et al., 2016), which is, according to EVT (Eccles et al., 1983), detrimental to their academic behavior and achievement (see **Chapters 1 and 2**). In this dissertation, which was based on EVT and data from the MoMa project (see **3.2**), investigation was made into how to help students understand the relevance of mathematics topics addressed in class (see **Chapter 3**). In Study 1 analysis was conducted of the relative effects of relevance-oriented teaching strategies and students' perception of their classmates' mathematics-related value beliefs on students' own mathematics-related value beliefs (see **Chapter 4**). In Study 2 the effects of the two MoMa relevance interventions in the classroom on students' mathematics-related competence beliefs, effort, and achievement were examined (see **Chapter 5**). Finally, in Study 3 exploration was made of the mechanisms through which a change in students' relevance beliefs may be triggered. This was done by analyzing students' responsiveness to the writing activities which were part of the MoMa relevance interventions (see **Chapter 6**). The central findings of these three studies will be summarized and discussed according to four major topics: (1) influences of the classroom context on students' value beliefs, (2) the effectiveness and mechanisms of the MoMa relevance interventions, (3) peers as a powerful source to communicate relevance, and (4) new insights into EVT. Subsequently, strengths and limitations of this dissertation and implications for future research will be explored. This chapter will close with concluding thoughts on the implications of the current work for educational policy and instructional practice.

# 7

# Discussion of the findings from the empirical studies

## 7.1 Influences of the classroom context on students' value beliefs

The focus of previous research framed within the Eccles et al. (1983) EVT primarily has been the last part of the model (cf., Figure 1), that is, students' competence beliefs and value beliefs and how they are connected with student outcomes (see e.g., review by Wigfield et al., 2009). Consequently, empirical evidence of how students' value development is influenced through important socializers within the context of the classroom (i.e., teachers and peers) has remained scarce. The results of some studies indicate that relevance-oriented teaching has the potential to motivate students, but in those studies not all value components were included and no comparison was made of the motivational potential of distinct instructional strategies (e.g., Assor et al., 2002; Lazarides & Rubach, 2017; Wang, 2012). Study 1 makes a unique contribution to research within the realm of EVT by including students' and teachers' perspective on the classroom context and by examining the influence of three relevance-oriented teaching strategies and students' perception of their classmates' mathematics-related value beliefs on the state and development of students' own mathematics-related value beliefs (see **2.1**).

The results of Study 1 differed according to the perspective on the classroom context, the value component under investigation, and the focus on either the state or the development of value beliefs. The student-reported strategy "stressing the practical applicability of mathematics" was shown to have a particularly strong influence on students' value development (increase in intrinsic, attainment, and utility values; decrease in cost). Cross-sectional associations were found between the teacher-reported strategy "demonstrating links between mathematics and other academic subjects" and some value beliefs of students; however, only the strategy "introducing new mathematics topics with everyday examples" contributed to a change in student motivation: This strategy helped keep ninth-graders from perceiving cost when doing mathematics (e.g., becoming anxious about learning mathematics). In addition, students' perception of their classmates' mathematics-related value beliefs predicted an increase in students' mathematics-related utility value. The results indicate that teaching mathematics with an emphasis on its practical applicability and using everyday examples when introducing new topics has the potential to motivate students over a long period of time. Furthermore, the class-specific value climate towards mathematics should be considered when trying to convey the relevance of mathematics topics. However, regarding the weak effects found especially for teacher-reported strategies, the results of Study 1 also imply that further research is needed to go more in depth concerning effective ways of conveying relevance within common instructional practices in mathematics (see **7.6**).

### 7.2 Effectiveness and mechanisms of the MoMa relevance interventions

While Study 1 was concerned with the question of how mathematics teachers convey relevance in their everyday instructional practices and the role students' classmates play in this context, Studies 2 and 3 were about the effectiveness of two scientific relevance interventions implemented in the classroom. Previous studies of American samples revealed the motivating potential of classroom-based interventions during which students identified the personal relevance of learning STEM topics (e.g., Hulleman & Harackiewicz, 2009, see **2.2.3**). Drawing on findings from such field studies (for a review, see e.g., Rosenzweig & Wigfield, 2016) and from laboratory-based experiments (e.g., on the interaction between competence beliefs and utility value beliefs; Durik et al., 2015), this dissertation aimed at comparing the effectiveness of two relevance intervention approaches with a German sample (MoMa project, see **3.2**). Both interventions consisted of two parts: (a) a classroom presentation with a confidence booster and examples of the usefulness of mathematics for students' current and future lives and (b) an individual writing assignment which differed according to condition. In a writing assignment based on social learning theories (e.g., Bandura, 1977b), students in the "quotations" condition commented on young adults' arguments about the relevance of mathematics; in a writing assignment adapted from prior classroom interventions (e.g., Hulleman & Harackiewicz, 2009), students in the "text" condition generated their own arguments about the relevance of mathematics. The interventions were implemented at the class level in a randomized controlled study design. Studies 2 and 3 add to previous research on classroom-based relevance interventions not only by broadly evaluating the effectiveness of the MoMa interventions but also by investigating the mechanisms underlying the intervention effects.

#### 7.2.1 Fostering students' motivation, behavior, and achievement through quotations conveying the relevance of mathematics

Previous classroom-based relevance intervention studies focused mainly on students' value beliefs and interest as short-term outcomes and grades as long-term outcomes (Hulleman et al., 2010; Hulleman & Harackiewicz, 2009; Woolley et al., 2013). To learn more about the breadth and sustainability of the effectiveness of relevance interventions, in Study 2 motivational, behavioral, and achievement outcomes were included at two points in time after the implementation of the interventions. Both subject- and task-specific competence beliefs (mathematics-related self-concept and homework-related self-efficacy) were investigated, as well as effort and test scores, thus covering a broad range of previously neglected outcomes. Results of prior research within the MoMa project indicated that the quotations condition had stronger and broader effects on students' short- and long-term value beliefs than the text condition (Gaspard, Dicke, Flunger, Brisson et al., 2015). In line with this, results of Study 2 indicate that the quotations condition promoted students' mathematics-related self-concept, homework-related self-efficacy, effort, and achievement in the short term and long term, whereas the text condition fostered students' long-term homework-related self-efficacy but no other variables under investigation. A summary of the main effects of the MoMa interventions is provided in Table 1.

Together with research by Gaspard et al. (2015; 2016), Study 2 thus gives a unique account of the potential of a newly developed intervention combining researcher-communicated utility value (in-class presentation) and peer-communicated utility value (writing task with statements from secondary school graduates, college students, etc.) to foster long-term outcomes over and above the target variable: students' utility value beliefs. Most interestingly, the quotations condition "outperformed" a condition adapted from an established relevance intervention approach (e.g., Hulleman & Harackiewicz, 2009) in its effectiveness: The effects of the text condition were less broad, less sustained, and—as shown in the following paragraph—depended more on the degree of students' responsiveness to the writing activity. As recent classroom interventions requiring students to generate arguments in writing about the relevance of a topic were not as effective as expected (Husman et al., 2017; Karabenick et al., 2017), future researchers developing relevance interventions could try to focus more on approaches similar to the MoMa quotations condition (see **7.6**).

Table 1

*Summary of significant main effects[a] of the MoMa relevance interventions*

| | Quotations | | Text | |
|---|---|---|---|---|
| | *ß* T2 | *ß* T3 | *ß* T2 | *ß* T3 |
| Mathematics-related value beliefs[b] | | | | |
| Utility value | .30 | .29 | .14 | .16 |
| Attainment value | .12 | .15 | – | – |
| Intrinsic value | *.08* | .14 | – | – |
| Cost | *-.08* | – | – | – |
| Mathematics-related competence beliefs[c] | | | | |
| Self-concept | .10 | *.09* | – | – |
| Homework self-efficacy | .16 | .20 | *.08* | .16 |
| Mathematics-related effort[c, d] | | | | |
| Teacher-rated effort | .14 | .12 | – | – |
| Self-reported effort | .13 | – | – | – |
| Achievement in mathematics[c] | | | | |
| Test scores | n/a | .18 | n/a | – |

*Notes.* [a] The coefficients are not fully comparable across publications due to their dependency on the modeling strategy and covariates used in the analyses.
[b] Gaspard et al., 2015; [c] Brisson et al., 2017; [d] Gaspard et al., 2017;
*ß* = standardized regression coefficient; T2 = six weeks after the intervention;
T3 = five months after the intervention; $p < .05$ (one-tailed $p$-value); coefficients in italics $p < .10$ (one-tailed $p$-value); – = nonsignificant effect; n/a = not applicable.

### 7.2.2 Towards a more refined knowledge about the intervention processes triggering changes in relevance beliefs

Assessing and analyzing students' responsiveness to the intervention activities is one way to learn about how relevance interventions work and why they work for some students but not for others (e.g., Murrah et al., 2017). Prior research has identified the quality and quantity of students' personal connections with the learning material and cognitive involvement in the task as central elements contributing to the effectiveness of relevance interventions (e.g., Harackiewicz et al., 2016; Hulleman & Cordray, 2009, see **2.2.4**). The aim of Study 3 was to shed more light on the differences in the effectiveness of the two MoMa intervention conditions. In Study 3 descriptive and causal analytical approaches were combined to determine whether the degree of students' responsiveness to the writing assignment (measured through the indicators "positive argumentation", "personal connections", and "in-depth reflections") affected the intervention effects on students' utility value beliefs. In addition, the characteristics of highly responsive and minimally responsive students were identified.

Overall, students' responsiveness to the writing activities was high in both intervention conditions. Furthermore, the intervention effects on students' utility value beliefs differed between responsive and nonresponsive students. The size of this difference depended on the degree of students' responsiveness to the assignment and on the condition. In the quotations condition, the most responsive students (i.e., those who had the highest values on all indicators) reported the most positive utility value beliefs after the intervention. However, even the least responsive students perceived mathematics to be more useful after the intervention than comparable students in the control group. In the text condition, utility value beliefs were improved for students with higher levels of responsiveness, whereas the least responsive students (i.e., a small minority of students who argued mainly against the usefulness of learning mathematics in their essays) did not profit from the intervention: On the contrary, their utility value beliefs were significantly lower after the intervention than those of comparable students in the control group. The students who were minimally responsive tended to be male, to lack conscientiousness, and to have negative mathematics-related utility value beliefs prior to the intervention. The results of Study 3 thus not only indicate that in the text condition the quality of students' relevance essays was more important for the intervention effects than in the quotations condition. They also suggest that boys, unconscientious students, and students with negative initial utility value beliefs should receive particular attention when implementing relevance interventions involving activities similar to that of the text condition in Study 3.

## 7.3 Peers: an effective source to communicate relevance

Theories and research have acknowledged growing peer-orientation of students throughout secondary school (e.g., Eccles, Midgley et al., 1993). The importance of peers as role models for students has been underlined for various academic outcomes, for example, engagement (e.g., Juvonen et al., 2012), self-efficacy and self-regulation (e.g., Schunk & Zimmerman, 2007), and achievement (e.g., Wentzel, 2005). In addition, secondary school students have been found to adapt increasingly their level of academic motivation to that of their peers (Kindermann et al., 1996; Kindermann, 2007). However, the salience of students' peers for students' competence and value development as defined in EVT has been neglected in the literature (see **2.1.3**). The results of the current dissertation provide new insights regarding this research gap.

In Study 1 investigation was made into the importance of students' classmates for students' value development. Results of the cross-sectional analyses of Study 1 indicate positive associations between students' perception of their classmates' mathematics-related value beliefs and students' mathematics-related intrinsic, attainment, and utility value beliefs. Negative associations have been found with the cost attributed to learning mathematics. Most interestingly, however, results of the longitudinal analyses indicate that peers influence students' perceptions of the usefulness of mathematical skills over time. Results of Study 1 thus indicate that reinforcing motivational processes as reported, for example, by Kindermann et al. (1996), also may occur in mathematics class with regard to students' perception of the relevance of mathematics.

Heightened peer orientation during adolescence also is reflected in the results of Studies 2 and 3, in which investigation was made into the effectiveness of the MoMa relevance interventions. Students who watched the MoMa presentation and then commented on statements made by slightly older peers (e.g., secondary school graduates, college students) about situations in which they needed mathematics profited enormously from the intervention: They were more confident in their mathematical skills, made more effort in mathematics class, and obtained higher scores on a mathematics test compared to students in the control group (Study 2). Even if the quality of students' written comments about their peers' statements was low, they found mathematics more relevant after the intervention than comparable students in the control group (Study 3). In contrast, the effectiveness of the text condition—which did not include any relevance statements made by peers—was found to be weaker and to depend more on the quality of students' responses to the writing task. It thus seems that the reading part containing peer-communicated relevance statements contributed to the compelling success of the quotations approach.

The results of Studies 1, 2, and 3 thus underline the importance of peers for students' motivation, in particular, for their utility value beliefs. When taking a closer look in Study 1 at the characteristics of the "peers", same-age and familiar peers (i.e., students' well known real classmates) influenced students' perception of the relevance of mathematics through actions indicating their mathematics-related value beliefs. Studies 2 and 3 revealed that value-related statements of peers also are important for students' motivation, behavior, and achievement when these peers are slightly older and entirely unfamiliar to the student. Interestingly, whether or not students are considered role models by same-age or younger students has been found to depend on the content of the information (see review of role model-observer similarity by Schunk, 1987): In the school context, information gained from same-age peers is particularly important when its content is highly valued. In contrast, older peers (or adults) may be particularly influential in issues where same-age peers generally are considered to be less knowledgeable. In fact, relevance interventions involving activities where both same-age and older peers serve as role models might be particularly effective in improving students' motivational development (see **7.6**).

### 7.4 New insights into EVT: relationship between competence beliefs and value beliefs

Results of correlational research indicate that competence beliefs and value beliefs are interrelated (e.g., Eccles et al., 2004; Nagengast et al., 2011) and that their association grows stronger over time (e.g., Jacobs et al., 2002; Trautwein et al., 2012). Until now, researchers have tended to assume that students value tasks because they are good at them (e.g., Jacobs et al., 2002; Wigfield et al., 2009). Yet, empirical support for this assumption is scarce, as experimental studies are needed to make descriptions of causes and effects (Shadish et al., 2002). Though less adopted, the assumption that students learn to become more confident in academic tasks and subjects because they value them also lacks empirical support from experimental studies (see **1.3.2**). There is first experimental evidence of the positive influence of students' utility value beliefs on students' competence beliefs as measured through students' outcome expectations (Hulleman et al., 2017). Yet, the nature of the relationship between students' competence beliefs and value beliefs needs to be investigated more comprehensively by distinguishing clearly between students' domain-specific and task-specific competence beliefs and different components of value.

This dissertation makes a substantial contribution to this gap in research. Results of Study 2 show that relevance interventions can have the power to raise students' subject-related competence beliefs and their task-related competence beliefs. More precisely, the quotations condition improved students' mathematics-related self-concept and homework-related self-efficacy. Though less effective overall, the text condition promoted students' mathematics homework-related self-efficacy. An additional analysis on the MoMa data reported in the Appendix (Part G) indicates that in both conditions, it seems likely that intervention effects on students' homework-related self-efficacy five months after the intervention were partially mediated through an increase in students' utility value beliefs six weeks after the intervention. The results of Study 2 and of the additional mediation analysis support the hypothesis that changes in students' utility value beliefs lead to changes in students' competence beliefs—in particular, task-specific self-efficacy beliefs—and thereby provide unique insight into the nature of the relationship between competence beliefs and value beliefs.

### 7.5 Strengths and limitations

In this dissertation investigation was made into how relevance of mathematics can be conveyed through conventional teaching approaches and through scientific interventions in a comprehensive way. The major strengths of this dissertation were the notable sample size of almost 2000 students, and the sound and up-to-date research design which involved several measurement points and appropriate analytical strategies to investigate the research questions. Furthermore, comprehensive instruments were used to assess both students' value beliefs (intrinsic, attainment, utility value, and cost) and students' competence beliefs (at the subject level and at the task level) in a differentiated way. In the following paragraph further strengths of this research are outlined and a few limitations to the research that should be kept in mind when interpreting the results of the current investigations are pointed out.

The choice of sample (ninth-grade students in Germany) and subject (mathematics) was valuable for two reasons. First, research on scientific relevance interventions had been confined to American samples and STEM-related subjects such as biology, science, or psychology (see overview in Table 1 of Study 2). Until now, it has been unclear whether similar relevance interventions would be effective in culturally different samples, for example with students in Germany, and in other STEM-related subjects such as mathematics (cf., Schmiedek, 2016, on the importance of culturally broad replications of psychological interventions). Second, 15-year-olds in Germany (i.e., ninth-grade students) have been found to be particularly at risk of attributing little utility value to mathematics (Reiss et al., 2016). At the same time, the societal and economic demand for mathematical skills is high in Germany and other Western countries (Institut der deutschen Wirtschaft, 2017; Renn et al., 2012), which is why it is important to find effective educational approaches to reverse or halt the downwards trend in students' motivation in mathematics. In academic-track schools, 15-year-olds are still a few years from graduating and thus may not yet have a clear idea of their future educational or career pathways, or know the extent to which mathematics will play a role in their life after graduation. Accordingly, the MoMa relevance interventions were designed particularly for the needs of students of this age group and school track. However, given this particular context, implementing the MoMa relevance interventions with other samples (e.g., students in vocational-track schools, college students, etc.) or in other subjects would require adapting the communicated relevance information to the respective target group and academic domain.

Second, as both teachers and students can be considered experts within the classroom context (Kunter & Baumert, 2006), the perspectives of both were included in the investigations made in Studies 1 and 2. Teacher reports on instructional practices are particularly valuable when estimations about a teaching strategy or method require professional subject-specific knowledge (Kunter & Baumert, 2006). Accordingly, in Study 1, teacher reports on specific instructional strategies (introducing new mathematics topics with everyday examples, demonstrating links with other academic subjects) were combined with student reports on a more general aspect of relevance-oriented mathematics instruction (stressing the practical applicability of mathematics). In Study 2, teachers' impression of the impact of the relevance interventions was included by having them rate their students' effort, whereas less visible, cognitive outcomes of the interventions (i.e., competence beliefs) were assessed through students' self-reports. The combination of these two important perspectives in Studies 1 and 2 constitutes a major strength of the current dissertation. Yet, data on the classroom context gained from teachers and students are subject to bias (resulting from teaching ideals, personal preferences, etc.; e.g., Pianta & Hamre, 2009)—especially when they are gathered via questionnaires in a retrospective way (e.g., Fahrenberg, Myrtek, Pawlik, & Perrez, 2007). Other research methods such as classroom observation and experience sampling could help to get closer to what actually happens in the classroom (see **7.6**).

Lastly, though in part adapted from previous relevance interventions (e.g., Hulleman & Harackiewicz, 2009), the content of MoMa interventions was newly developed and the study design highly innovative compared to previous classroom-based relevance interventions (e.g., Table 1 in Study 2). Innovative features included, for example, the integration of a short competence booster (cf., Durik et al., 2015), the combination of communicated and self-generated relevance information (cf., Canning & Harackiewicz, 2015), and, in particular, the inclusion of peer-communicated relevance information in the quotations condition (cf., motivational importance of peers, e.g., Study 1; Kindermann, 2007). Through implementation at the class level the interventions were highly adapted to students' genuine learning context and, therefore, the risk of diffusion effects was reduced (Craven et al., 2001). The interventions were implemented by researchers in a highly standardized way, which was important for analysis and comparison of the motivational potential of the newly developed intervention activities. Indeed, results of Studies 2 and 3 revealed differences in the effectiveness and related mechanisms of the two intervention conditions, thereby indicating which approach could be more worthwhile implementing on a larger scale in educational practice. However, the current results do not allow for conclusions to be drawn regarding the effectiveness of the MoMa interventions when delivered by teachers—or even older peers (see **7.6**).

### 7.6 Implications for future research

This dissertation provides valuable new insights into how the relevance of mathematics can be conveyed in the classroom. Future researchers are invited to build upon this work by further exploring students' competence beliefs and value beliefs in the classroom. Most centrally, replicating the current studies with different samples and for different academic subjects, and including long-term outcomes (e.g., motivational beliefs across school years, course choices, career choices, etc.) would be necessary to gain insight into the generalizability of the findings and the long-term impact of relevance-oriented teaching and interventions.

In future, researchers could consider using additional instruments and variables to investigate ways to convey relevance in everyday instructional practice. The inclusion of both teachers' and students' reports on relevance-oriented teaching strategies would be helpful to disentangle the relative importance of different views on mathematics instruction for students' value development. Observer ratings, experience sampling data, and qualitative measures (e.g., coding mathematics exercise sheets) might be helpful to assess relevance-oriented instruction in a less biased way than by using questionnaires (e.g., Fahrenberg et al., 2007). Furthermore, interviews with teachers have revealed that it might be difficult to convey the direct relevance of some mathematics topics and tasks to students' lives (Turner et al., 2011). Therefore, it would be interesting to examine in addition the influence of emphasizing the general value of learning for students' current and future value development in mathematics (see also Brophy, 1999; Wentzel & Brophy, 2014).

Second, the results of the studies conducted for this dissertation indicate that assessing students' competence beliefs and value beliefs in a comprehensive and differentiated way is crucial to understand better the nature and development of, and relationship between, students' competence beliefs and value beliefs. More precisely, the effects found in Studies 1 and 2 differed according to the value component (intrinsic, attainment, utility, or cost) and type of competence belief (domain- or task-specific) investigated. By differentiating the four value components also in the future, inconsistencies in the measurement of students' academic values could eventually be overcome (see **1.3.1**) and more detailed knowledge about contextual influences on single value components could be gained. Similarly, future research framed in EVT could profit from distinguishing students' domain- and task-specific competence beliefs as distinct outcomes of relevance interventions. Beyond that, to understand better the directional influences between students' competence beliefs and students' value beliefs, competence beliefs and value beliefs should both be assessed in a differentiated way when either of them constitutes the target of an intervention. Despite the huge number of competence experiments, intervention effects have not yet been reported on all components of students' value beliefs simultaneously (for reviews, see Haney & Durlak, 1998; O'Mara et al., 2006).

Third, to explore further how relevance can be conveyed successfully in the classroom through short scientific interventions, the quotations condition could be developed further and compared with previous approaches. Instead of using a writing activity to personalize the message of the intervention (cf., Yeager & Walton, 2011), researchers could, for example, develop partner activities where students first summarize previously obtained utility information (e.g., from statements made by older peers) and then explain the personal usefulness of mathematics to their partners. Combining such replication and production tasks as partner activities is an effective way to promote meaningful knowledge structures (cf., research on generative learning strategies, e.g., Fiorella & Mayer, 2016). In addition, following SDT (Ryan & Deci, 2000), meta-analyses of cooperative learning (e.g., Kyndt et al., 2013; Lou et al., 1996), and findings on increasing peer-orientation during adolescence (e.g., Kindermann et al., 1996), such partner activities could enhance students' motivation to engage in the task and, in turn, enhance students' responsiveness to the intervention.

Fourth, assessing and investigating intervention fidelity (e.g., students' responsiveness to intervention tasks) should be standard in classroom-based experimental research. Until now, students' responsiveness to relevance interventions has been assessed primarily to examine why an intervention did not produce any effects (Husman et al., 2017; Karabenick et al., 2017) or produced weaker effects than a corresponding laboratory-based intervention (Hulleman &

Cordray, 2009). However, to learn about the mechanisms leading to effects, intervention fidelity also must be studied in effective intervention programs. Causal effects analyses (e.g., CACE models, Jo, 2002) are crucial to determine what happens if students do not complete intervention activities as intended (e.g., by arguing against the usefulness of mathematics), and to characterize these students to better meet their motivational needs in the future. Furthermore, in future interventions, researchers could assess students' responsiveness not only to the part involving the generation of utility information (e.g., writing essays), but also to activities aimed at communicating utility information (e.g., listening to a presentation, reading quotations). Individual factors such as emotional states or cognitive activation, which might correspond with students' engagement in the intervention activities, could be assessed, for example, through computerized experience sampling methods (e.g., Pekrun & Linnenbrink-Garcia, 2012).

Fifth, the importance of various modes and sources of communicating utility information to students during relevance interventions could be disentangled by conducting laboratory experiments. The quotations condition included utility information communicated by an adult (listening activity) or communicated by peers (reading activity). Combining activities appealing to different channels (auditory, visual) might have contributed to the effectiveness of the quotations condition (cf., research on multimedia learning and memory processing, e.g., Mayer & Moreno, 1998). In future research, the mode of communicating utility value could be varied by creating different experimental conditions including live presentations, audio recordings, or video material vs. a condition involving written material (or combined conditions). Similarly, experimental studies using varying sources of utility information such as same-age peers, slightly older peers, and adults could provide further insight into the most effective modes and sources of communicated utility value (cf., Hoogerheide, van Wermeskerken, Loyens, & van Gog, 2016, who compared the effectiveness of identical explanations when delivered by peers to when delivered by adults for learning science).

Sixth, to pave the way for relevance interventions to enter educational practice, future research is needed in which teachers are included in the implementation process of relevance intervention programs. The effectiveness of the MoMa interventions could be tested and compared when implemented in the same standardized way by teachers as opposed to researchers. However, following a design-based research approach (e.g., Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003), the MoMa intervention material also could be refined in close collaboration between scientists and teachers using several cycles of implementation, evaluation, and individual adaptation. During these cycles, teachers or researchers might propose new intervention elements, distribute elements of the 90-minute MoMa program over several lessons, and thus reduce the standardization of the intervention material. This is how, on the one hand, teachers can bring in their own ideas of how to integrate researcher-developed intervention activities in everyday instructional practice (cf., importance of teachers' beliefs for instructional behaviors, e.g., Reeve et al., 2014; Turner et al., 2011). On the other hand, in such collaboration, teachers also could be informed about how to enhance students' competence beliefs and value beliefs on a regular basis, for example, by regularly using relevance-oriented teaching strategies and being sensitive to group dynamics regarding mathematics-related value beliefs (cf., Study 1; see also Woolley et al., 2013). Teachers also could be trained on how to enhance the overall motivational quality of their teaching by including, for example, autonomy support or high structuredness—which might even reinforce the intervention effects (cf., Jang et al., 2010; Lazarides & Ittel, 2012; see also **2.1**). As a result, the effectiveness of different implementations could be compared: (a) standardized vs. adapted use of relevance intervention material by teachers, (b) one-time vs. continuous integration of relevance information in teachers' mathematics instruction, and (c) providing relevance information without vs. within an overall autonomy-supportive and/or structured mathematics teaching style.

Lastly, further research is needed on how students can be encouraged to support each other's mathematics-related value beliefs on a regular basis. In fact, tutoring programs in which students help classmates or younger fellows in learning activities (Topping, 2000) have been shown to improve students' achievement, attitudes towards school, and classroom behavior. Interestingly in the subject of mathematics, these effects were found for both tutors (i.e., students assisting others) and tutees (i.e., students receiving help), with slightly stronger effects for reciprocal than for unidirectional tutoring (for reviews, see e.g., Ramani, Zippert, & Daubert, 2016; Robinson, Schofield, & Steers-Wentzell, 2005).These effects often have been framed within role theory assuming that tutors adopt behaviors and attitudes consistent with the role identity of a teacher (e.g., conveying new information, using verbal reinforcements, liking the subjects taught, valuing school in general; e.g., Allen & Feldman, 1976; Sarbin, 1976; Turner, 2006). Due to statutory equality, tutees also may identify more with peer tutors than with teachers; in turn, tutors who are aware that they act as role models may show more socially desired academic behaviors (Allen & Feldman, 1976). In previous peer tutoring programs, tutors rarely have been encouraged to actively transmit the norms and values of learning mathematics to their tutees, which might be very important in adolescence (cf., Ramani et al., 2016). Thus, in future research analysis could be made of whether peer tutoring programs in which tutors are advised to consciously convey information on the relevance of mathematics to their tutees has the power to improve students' mathematics-related value beliefs and value-related behaviors.

## 7.7 Implications for educational policy and practice

Using specific teaching strategies or a researcher-led intervention program, the current research describes how and to what extent students' perceptions of the relevance and value of mathematics were fostered in the classroom context. By nature, the current findings reflect what worked well and what worked less well with a specific sample (students in Grade 9, academic track schools in Germany) in a specific subject (mathematics) and in a specific study design—and thus cannot tell what works or will work in everyday instructional practice. Consequently, the current findings can hardly be translated into general guidelines or clear rules for action;

however, they can help politicians and teachers change perspectives on teaching and learning mathematics (cf., Biesta, 2007).

In particular, the findings of this dissertation provide further support for the importance of students' perceptions of the relevance of mathematics as a core motivational factor (Eccles et al., 1983). Changes in students' perception of the utility value of mathematics have been shown to correlate with changes in students' mathematics- and homework-related competence beliefs, the effort they make in mathematics-related classes, and even their test performance in mathematics. Policy makers and teachers should be aware of the importance of students' utility value beliefs and should know that they are highly malleable and can even be improved during very short researcher-led interventions which are easy to implement in the classroom. Regarding the low mean ratings of the usefulness of mathematics reported by secondary school students in Germany compared to those of their peers learning mathematics in other OECD countries (Reiss et al., 2016) and the actual importance of mathematical skills for private life (Renn et al., 2012) and in the job market (Institut der deutschen Wirtschaft, 2017), it seems that the relevance of mathematical knowledge and skills is underemphasized in mathematics instruction in Germany. It would be desirable if relevance-oriented mathematics instruction could be part of teacher training and the standard mathematics curriculum. Introducing new mathematics topics with everyday examples or providing examples of situations outside school or after schooling in which mathematical skills can be applied may be part of a general teaching approach with an emphasis on the practical value of mathematics.

However, the huge challenge of maintaining students' motivation in mathematics throughout secondary school cannot be done by teachers alone who integrate relevance information into their mathematics instruction—students also need peers as positive role models at school (and, ideally, positive role models outside school, e.g., parents, cf., Häfner et al., 2017; Harackiewicz et al., 2010). Classmates' behaviors indicating how much they personally value mathematics and older peers' reports on where they need mathematical skills in their personal lives affect students' own motivation to learn mathematics. It would be desirable if teachers were sensitive to motivational dynamics within the class as a group. At the same time, peers seem to be a valuable source when aiming at communicating the relevance of mathematics to students. Accordingly, students should be made aware that they may act as motivational role models for other students, particularly younger students. Older students including those in tertiary education could even be encouraged to participate with younger students in cross-age peer tutoring projects on the value of learning mathematics.

In summary, to halt or even reverse the downward trend in students' motivation in mathematics, holistic approaches are needed which involve both the teacher and peers conveying relevance of mathematics topics addressed in class. Professional learning partnerships between schools and researchers, who jointly develop, implement, and evaluate programs that promote student motivation by including peers as a source of relevance information might be key for integrating findings from scientific relevance interventions into instructional practice.

## References

Allen, V. L., & Feldman, R. S. (1976). Studies on the role of tutor. In V. L. Allen (Ed.), Children as teachers. Theory and research on tutoring (pp. 113–130). New York: Academic Press.

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. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1), 1–22. https://doi.org/10.1111/j.1741-5446.2006.00241.x

Brophy, J. E. (1999). Teaching. Lausanne, Switzerland: PCL.

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility value information. *Motivation Science, 1*(1), 47–71. https://doi.org/10.1037/mot0000015

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Craven, R. G., Marsh, H. W., Debus, R. L., & Jayasinghe, U. (2001). Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. *Journal of Educational Psychology, 93*(3), 639–645. https://doi.org/10.1037/0022-0663.93.3.639

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

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). New York: Freeman.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Maciver, D. (1993). Development during adolescence—The impact of stage-environment fit on young adolescents experiences in schools and in families. *American Psychologist, 48*(2), 90–101. https://doi.org/10.1037//0003-066x.48.2.90

Eccles, J. S., Vida, M. N., & Barber, B. (2004). The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *Journal of Early Adolescence, 24*(1), 63–77. https://doi.org/10.1177/0272431603260919

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment: Monitoring behavior in daily life settings. *European Journal of Psychological Assessment, 23*(4), 206–213. https://doi.org/10.1027/1015-5759.23.4.206

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Trautwein, U., & Nagengast, B. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*(9), 1226–1240. https://doi.org/10.1037 /dev0000028

Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B. M., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open, 2*(2), 1–14. https://doi.org /10.1177 /2332858416649168

Häfner, I., Flunger, B., Dicke, A.-L., Gaspard, H., Brisson, B. M., Nagengast, B., & Trautwein, U. (2017). Robin Hood effects on motivation in math: Family interest moderates the effects of relevance interventions. *Developmental Psychology, 53*(8), 1522–1539. https://doi.org/ 10.1037/dev0000337

Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytical review. *Journal of Clinical Child Psychology, 27*(4), 423–433.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*(5), 745–765. https://doi.org/10.1037 /pspp0000075

Harackiewicz, J. M., Hulleman, C. S., Rozek, C., Katz-Wise, S. L., & Hyde, J. S. (2010, March). *Parents' understanding of the utility value of STEM courses for high school students.* Paper presented at the 13th Biennial SRA Conference. Philadelphia, PA.

Hoogerheide, V., van Wermeskerken, M., Loyens, S. M. M., & van Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction, 44*, 22–30. https://doi.org/10.1016/j.learninstruc.2016.02.004

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*(1), 88–110. https://doi.org/10.1080/1934574080253925

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*(4), 880–895. https://doi.org/10.1037/A0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*(3), 387–404. https://doi.org/10.1037/edu0000146

Husman, J., Nelson, K., & Cheng, K. (2017, August). *Intervening to promote engineering student success in technical, nonmajor, required barrier courses.* Poster presented at the 17th Biennial EARLI Conference. Tampere, Finland.

Institut der deutschen Wirtschaft. (2017). MINT-Frühjahrsreport 2017: *MINT-Bildung: Wachstum für die Wirtschaft, Chancen für den Einzelnen [Education in STEM: Economic growth, opportunities for the individual]*. http://www.mintzukunftschaffen.de

/uploads/media/FINAL_MINT-Fruehjahrsreport_2017.pdf

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*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588–600. https://doi.org/10.1037/a0019682

Jo, B. (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics, 27*(4), 385–409. https://doi.org/10.3102/10769986027004385

Juvonen, J., Espinoza, G., & Knifsend, C. (2012). The role of peer relationships in student academic and extracurricular engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 387–492). New York: Springer.

Karabenick, S. A., Albrecht, J., & Rausch, N. (2017, August). *Examining the self-generation of subjective task value: A mixed-methods investigation*. Paper presented at the 17th Biennial EARLI Conference. Tampere, Finland.

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Development, 78*(4), 1186–1203. https://doi.org/10.1111/j.1467-8624.2007.01060.x

Kindermann, T. A., McCollam, T. L., & Gibson, E. (1996). Peer networks and students' classroom engagement during childhood and adolescence. In J. Juvonen & K. R. Wentzel (Eds.), *Social* *motivation: Understanding children's school adjustment* (pp. 279–312). Cambridge, England: Cambridge University Press.

Klein, F., & Schimmack, R. (1907). *Vorträge über den Mathematischen Unterricht an den Höheren Schulen. Mathematische Vorlesungen an der Universität Göttingen.* Leipzig: B. G. Teubner.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231–251. https://doi.org/10.1007/s10984-006-9015-7

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning: Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133–149. https://doi.org/10.1016/j.edurev.2013.02.002

Lazarides, R., & Ittel, A. (2012). Instructional quality and attitudes toward mathematics: Do self-concept and interest differ across students' patterns of perceived instructional quality in mathematics classrooms? *Child Development Research, 2012*, 1–11. https://doi.org/10.1155/2012/813920

Lazarides, R., & Rubach, C. (2017). Instructional characteristics in mathematics classrooms: Relationships to achievement goal orientation and student engagement. *Mathematics Education Research Journal, 29*(2), 201–217. https://doi.org/10.1007/s13394-017-0196-4

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423–458. https://doi.org/10.3102/00346543066004423

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312–320. https://doi.org/10.1037/0022-0663.90.2.312

Murrah, W. M., Kosovich, J. J., & Hulleman, C. S. (2017). A framework for incorporating intervention fidelity in educational evaluation studies. In G. Roberts, S. Vaughn, T. Beretvas, & V. Wong (Eds.), *Treatment fidelity in studies of educational intervention* (pp. 39–60). New York: Routledge.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*(8), 1058–1066. https://doi.org/10.1177/0956797611415540

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181–206. https://doi.org/10.1207/s15326985ep4103_4

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). New York: Springer.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Ramani, G. B., Zippert, E., & Daubert, E. (2016). The influence of same- and cross-age peers on children's literacy and mathematical development. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of social influences in school contexts. Social-emotional, motivation, and cognitive outcomes* (pp. 96–112). New York, London: Routledge Taylor & Francis Group.

Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H.,. . . Wang, C. K. J. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion, 38*(1), 93–110. https://doi.org/10.1007/s11031-013-9367-0

Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E., & Köller, O. (Eds.). (2016). *PISA 2015: Eine Studie zwischen Kontinuität und Innovation.* Münster: Waxmann. Retrieved from http://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtband_und_Zusammenfassung_2012/PISA_2015_eBook.pdf

Renn, O., Duddeck, H., Menzel, R., Holtfrerich, C.-L., Luca, K., Fischer, W.,. . . Pfenning, U. (2012). *Stellungnahmen und Empfehlungen zur MINT-Bildung in Deutschland auf der Basis einer europäischen Vergleichsstudie*. Berlin.

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*(4), 327–362. https://doi.org/10.1007/s10648-005-8137-2

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM Motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist, 51*(2), 146–163. https://doi.org/10.1080/00461520.2016.1154792

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78.

Sarbin, T. R. (1976). Cross-age tutoring and social identity. In V. L. Allen (Ed.), *Children as teachers. Theory and research on tutoring* (pp. 27–40). New York: Academic Press.

Schmiedek, F. (2016). Methods and Designs. In T. Strobach & J. Karbach (Eds.), *Cognitive training. An overview of features and applications* (pp. 9–18). Cham: Springer International.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*(2), 149–174. https://doi.org/10.2307/1170234

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly, 23*(1), 7–25.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Belmont, CA: Wadsworth Cengage Learning.

Topping, K. (2000). *Tutoring (5th ed.). Educational Practice Series*. Lausanne, Switzerland: PCL.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. https://doi.org/10.1037/A0027470

Turner, J. C., Warzon, K. B., & Christensen, A. (2011). Motivating mathematics learning: Changes in teachers' practices and beliefs during a nine-month collaboration. *American Educational Research Journal, 48*(3), 718–762. https://doi.org/10.3102/0002831210385103

Turner, R. H. (2006). Role theory. In J. H. Turner (Ed.), *Handbooks of sociology and social research: Handbook of sociological theory* (pp. 233–254). Dordrecht: Springer.

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology, 48*(6), 1643–1657. https://doi.org/10.1037/A0027247

Wentzel, K. R. (2005). Peer relationships, motivation, and academic performance at school. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 279–296). New York and London: Guilford.

Wentzel, K. R., & Brophy, J. E. (2014). *Motivating students to learn (4th ed.).* New York: Routledge.

Wigfield, A., Tonks, S., & Klauda, S. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–75). New York and London: Routledge.

Woolley, M. E., Rose, R. A., Orthner, D. K., Akos, P. T., & Jones-Sanpei, H. (2013). Advancing academic achievement through career relevance in the middle grades: A longitudinal evaluation of CareerStart. *American Educational Research Journal, 50*(6), 1309–1335. https://doi.org/10.3102/0002831213488818

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301. https://doi.org/10.3102 /0034654311405999

# APPENDIX

## Part A) Intervention materials

### Quotations condition

**Sample quotations:**

"Logical thinking, the kind you learn in math, is a basic requirement in many professions, but when you are 15 or 16, you don't know that. Then, you think 'I don't ever want to have anything to do with math again!' And when you later want to study something specific, like psychology or economics, you suddenly lack the basics and you think 'Oh my God! I should have paid more attention before!' It's really hard to make up for what you missed earlier; you need to be really strong-willed." (Diana, 20 years, psychology student)

"To me, math is an important part of general education. Even though it's often abstract and theoretical, you need math skills to be able to talk about certain things. Like understanding statistics about climate change or cancer risks, for instance. Or to be able to deal with computers. You also need it to pass your driver's license exam, math really helps with that!" (Ibrahim, 23 years, chemistry student)

**Work assignments:**

1) Are there any statements you have heard by others or thought about yourself in a similar way before? If so, in what situation?

2) Evaluate the statements: Which ones can you relate to? What is it that convinces you about these statements?

3) Rank how important you personally find the quotations from least to most important (…) and explain your ranking in detail.

4) What is your most important take-away message why math is useful?

### Text condition

**Instruction:**

1) In this study, we would like to find out how students perceive the utility of math for their current and future lives. Therefore, we would like for you to tell us why math is important and useful in your personal life. Think about situations in your everyday life in which you will need math skills now or in the future, and think about why math skills may be important for your professional plans. Please do not make judgments about whether or not you like math. It is not about fun with or interest in math but about the personal utility of math. (…) The more reasons you can find, the more helpful it will be to us as we will gain better insight into your opinion. (…)

2) What is your most important take-away message why math is useful?

## Part B) Coding of students' relevance essays (examples)

### Quotations condition

**Student's responses to work assignments 2) and 3) (summarized):**

Quotation 1: You need logical thinking. Yet there are *many professions you don't need logical thinking for*. Still good argument.

Quotation 2: If you strive and *if you make an effort, you can do it*, you can understand math. True argument.

Quotation 3: *Useless statement*, bad argument.

Quotation 4: It is *important for general education*. True argument.

Quotation 5: You can be happy when you solve difficult problems, but this statement is not really true. You are *only happy when you solve very difficult math problems*.

Quotation 6: It is true that math is not a subject you have to learn by heart and that math *can be fun*. Good statement.

**Student's take-away message:**

**I** will often need it for **my** *future profession*, in this case, *paratrooper*. But in *other professions, you don't need* it often.

Table B1
*Sample coding scheme (quotations condition)*

| Indicator | Examples and counts of coded instances | | **Value** |
|---|---|---|---|
| Utility arguments (marked in *italics*) | positive arguments (*grey shade*) | negative arguments | |
| | - understandable subject<br>- important for general education<br>- math can be fun<br>- needed in future profession | - useless in many professions (2x)<br>- useless for everyday finances<br>- not happy when solving easy problems | |
| Count | 4 | 4 | **2** |
| Personal connections (marked by underlining) | first person pronouns (**bold**) | impersonal pronouns | |
| | - I<br>- my | - you (12x) | |
| Count | 2 | 12 | **2** |
| In-depth reflections (marked by *italics*) | self-generated arguments (*circled*) | arguments mentioned in presentation | |
| | - use of own formulations/ thoughts when referring to quotes 1, 2, 6<br>- reference to paratrooper | - formulations copied when referring to quotes 4, 5<br>- no new thoughts for quote 3 | |
| Count | 4 | 3 | **3** |

## Text condition

**Student's text:**

At the moment, math is most important for **me** at *school*. Getting a good grade is the most important thing for **me** at the moment to meet **my** parents' and **my own** *expectations*. Besides, math helps **me** immensely to *collect money* for the [name of local car magazine], and to serve and *bill* customers at events, not having to note down everything but simply calculating quickly in my head. Math also helps **me** to *keep track of* **my own** *finances*, to know what **I** can afford, etc. What **I** find most important though is that good math skills give you so many *opportunities*, e.g. *applying for* different kinds of *study programs* or different kinds of *jobs*. In *everyday life*, math does not play a big role for **me** because it is *just part of it*, so you don't notice it much.

**Student's take-away message:**

Having many options for **my** future studies and professional life.

Table B2

*Sample coding scheme (text condition)*

| Indicator | Examples and counts of coded instances | | Value |
|---|---|---|---|
| Utility arguments (marked in *italics*) | positive arguments (*grey shade*) | negative arguments | |
| | - important at school (good grade, meet expectations)<br>- collect money for journal<br>- bill customers at events<br>- keep track of finances/ know what to afford<br>- apply for study programs<br>- apply for a job<br>- part of everyday life | | |
| Count | 7 | 0 | 3 |
| Personal connections (marked by underlining) | first person pronouns (**bold**) | impersonal pronouns | |
| | - me (5x)<br>- my (own) (4x)<br>- I (2x) | - you (2x) | |
| Count | 11 | 2 | 3 |
| In-depth reflections (marked by *italics*) | self-generated arguments (*circled*) | arguments mentioned in presentation | |
| | - important at school (good grade, meet expectations)<br>- collect money for journal<br>- bill customers at events<br>- keep track of finances/ know what to afford<br>- part of everyday life | - apply for study programs<br>- apply for a job | |
| Count | 5 | 2 | 3 |

## Part C) Intercorrelations of all variables under investigation

Table C1

*Correlations between students' intervention responsiveness, individual characteristics, and classroom perceptions at T1, as well as math-related utility value at T2 and T3 in the quotations condition (below diagonal) and in the text condition (above diagonal)*

| | | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | | (7) | | (8) | | (9) | | (10) | | (11) | | (12) | | (13) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* |
| (1) | Responsiveness | - | - | -.12 | .001 | .10 | .021 | .20 | .000 | .17 | .001 | .18 | .000 | .18 | .000 | .23 | .000 | .25 | .000 | .17 | .001 | -.13 | .020 | .31 | .000 | .24 | .000 |
| (2) | Gender (1 = male) | -.03 | .473 | - | - | -.03 | .461 | -.17 | .000 | .02 | .662 | .17 | .000 | .09 | .024 | .16 | .000 | .05 | .216 | .13 | .001 | -.07 | .083 | .01 | .743 | .08 | .036 |
| (3) | Cognitive ability | .04 | .139 | -.01 | .797 | - | - | .01 | .761 | .30 | .000 | .36 | .000 | .12 | .007 | .25 | .000 | .11 | .005 | .02 | .720 | -.05 | .348 | .07 | .076 | .02 | .698 |
| (4) | Conscientiousness | .17 | .001 | -.09 | .034 | -.11 | .013 | - | - | .16 | .000 | .21 | .000 | .22 | .000 | .20 | .000 | .25 | .000 | .10 | .025 | -.13 | .000 | .18 | .000 | .16 | .000 |
| (5) | Math test score | .20 | .000 | .12 | .006 | .25 | .000 | .10 | .054 | - | - | .51 | .000 | .27 | .000 | .42 | .000 | .18 | .000 | .04 | .518 | -.09 | .175 | .20 | .000 | .18 | .000 |
| (6) | Self-concept | .12 | .039 | .28 | .000 | .18 | .000 | .20 | .001 | .49 | .000 | - | - | .47 | .000 | .74 | .000 | .38 | .000 | .16 | .000 | -.11 | .003 | .31 | .000 | .23 | .000 |
| (7) | HW self-efficacy | .15 | .001 | .09 | .051 | .02 | .686 | .33 | .000 | .24 | .000 | .47 | .000 | - | - | .44 | .000 | .27 | .000 | .20 | .000 | -.13 | .016 | .26 | .000 | .24 | .000 |
| (8) | Intrinsic value | .21 | .000 | .14 | .002 | .12 | .003 | .29 | .000 | .43 | .000 | .71 | .000 | .45 | .000 | - | - | .50 | .000 | .36 | .000 | -.17 | .011 | .40 | .000 | .33 | .000 |
| (9) | Utility value | .20 | .001 | .12 | .006 | -.02 | .593 | .33 | .000 | .26 | .000 | .41 | .000 | .37 | .000 | .55 | .000 | - | - | .36 | .000 | -.17 | .009 | .65 | .000 | .56 | .000 |
| (10) | Class' math valuing | .14 | .005 | -.07 | .107 | -.01 | .837 | .09 | .120 | .11 | .056 | .17 | .002 | .22 | .000 | .30 | .000 | .29 | .000 | - | - | -.40 | .000 | .28 | .000 | .31 | .000 |
| (11) | Disruptions in class | -.07 | .185 | .04 | .337 | .04 | .467 | -.09 | .109 | -.06 | .162 | -.04 | .368 | -.06 | .215 | -.09 | .038 | -.12 | .005 | -.23 | .012 | - | - | -.19 | .003 | -.20 | .004 |
| (12) | Utility value T2 | .19 | .000 | .11 | .009 | -.05 | .197 | .25 | .000 | .17 | .000 | .29 | .000 | .29 | .000 | .42 | .000 | .70 | .000 | .22 | .000 | -.09 | .008 | - | - | .64 | .000 |
| (13) | Utility value T3 | .16 | .001 | .07 | .111 | -.01 | .761 | .20 | .000 | .17 | .000 | .27 | .000 | .18 | .000 | .34 | .000 | .59 | .000 | .16 | .000 | .00 | .984 | .67 | .000 | - | - |

*Notes*. T = time; HW = homework; *r* = Pearson's correlation coefficient; *p* = *p*-value.

**Part D) Sample syntax of the CACE models**
(Jo et al., 2008; Muthén & Muthén, 1998-2012)

```
Data: file = compliance.dat;

Variable: names are SID Class_ID Schul_ID intcond text quot cutoff4 cutoff5 cutoff6 cutoff7 cutoff8
  cutoff9 cutoff10 sumIR util2 util3 sex kft con score selfc hwse util intr clval disrup;

Usevariables are quot cutoff4 util2 util3 sex kft con score selfc hwse intr util clval disrup;

Classes = c(2);                                  ! analysis is done assuming 2 classes.
Subpopulation = (intcond==1 or intcond==3);      ! populations: quotations, control group.
Categorical = cutoff4;                           ! cutoff4: binary indicator of compliance.
Missing = all (-99);                             ! missing is coded -99.
Cluster = Class_ID;                              ! adjust SEs for nesting of data in classes.

Analysis: type = mixture missing complex; algorithm = integration; integration = montecarlo;

Model:
%overall%                                                    ! overall model.
[C#1];                                                       ! logit intercept.
C#1 on sex kft con score selfc hwse intr util clval disrup;  ! logit coefficient.
util2 on quot sex kft con score selfc hwse intr util clval disrup; ! effects on utility T2.
util3 on quot sex kft con score selfc hwse intr util clval disrup; ! effects on utility T3.
[util2 util3];                                               ! outcomes intercept.
util2 util3;                                                 ! outcomes residual variance.
util2 with util3;                                            ! correlation outcomes.
sex kft con score selfc hwse intr util clval disrup;         ! covariates residual variance.

%c#1%                                                        ! noncompliers.
[cutoff4$1@15];                                              ! probability of complier = 0.
util2 on quot (a);                                           ! intervention effect T2.
util3 on quot;                                               ! intervention effect T3.
util2 on quot sex kft con score selfc hwse intr util clval disrup; ! effect covariates on utility T2.
util3 on quot sex kft con score selfc hwse intr util clval disrup; ! effect covariates on utility T3.
[util2 util3];                                               ! outcomes intercepts.
util2 util3;                                                 ! outcomes residual variances.
util2 with util3;                                            ! correlation outcomes.

%c#2%                                                        ! compliers.
[cutoff4$1@-15];                                             ! probability of complier = 1.
util2 on quot (b);                                           ! intervention effect T2.
util3 on quot;                                               ! intervention effect T3.
util2 on quot sex kft con score selfc hwse intr util clval disrup; ! effect covariates on utility T2.
util3 on quot sex kft con score selfc hwse intr util clval disrup; ! effect covariates on utility T3.
[util2 util3];                                               ! outcomes intercepts.
util2 util3;                                                 ! outcomes residual variances.
util2 with util3;                                            ! correlation outcomes.

Model test:                                                  ! Wald-X²-test.
a=b;                                                         ! compare effects complier/NC.
```

# Part E) Full tables on the sensitivity CACE analyses

Table E1

*Intervention Effects on Students' Math Utility Value Beliefs in the Quotations Condition Depending on Students' Responsiveness: Cutoff on Responsiveness Index below Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 4 | | | | | | 5 | | | | | | 6 | | | | | |
| C min. value | 5 | | | | | | 6 | | | | | | 7 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 91 (8%) | | | | | | 136 (11%) | | | | | | 145 (12%) | | | | | |
| Complier | 1105 (92%) | | | | | | 1059 (89%) | | | | | | 1051 (88%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (*SE*) | *p* | $\beta_{NC}$ | (*SE*) | *p* | $\beta_{NC}$ | (*SE*) | *p* | $\beta_{NC}$ | (*SE*) | *p* | $\beta_{NC}$ | (*SE*) | *p* | $\beta_{NC}$ | (*SE*) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .35 | (.23) | .125 | .41 | (.32) | .199 | .35 | (.24) | .142 | .45 | (.25) | .071 | .38 | (.23) | .105 | .45 | (.25) | .077 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.02 | (.28) | .954 | -.08 | (.27) | .780 | .14 | (.24) | .557 | .17 | (.19) | .370 | .12 | (.22) | .578 | .17 | (.18) | .338 |
| Cognitive ability | -.41 | (.13) | .002 | -.29 | (.18) | .112 | -.23 | (.08) | .003 | -.09 | (.09) | .329 | -.22 | (.08) | .003 | -.08 | (.09) | .344 |
| Conscientiousness | .14 | (.09) | .129 | .09 | (.12) | .471 | .06 | (.09) | .503 | .01 | (.09) | .880 | .05 | (.09) | .527 | .02 | (.08) | .769 |
| Test score | .10 | (.15) | .491 | .17 | (.20) | .400 | .05 | (.10) | .631 | .04 | (.14) | .773 | .08 | (.09) | .421 | .05 | (.13) | .705 |
| Self-concept | .15 | (.13) | .229 | .09 | (.19) | .628 | .17 | (.12) | .175 | .17 | (.15) | .265 | .14 | (.11) | .182 | .15 | (.15) | .309 |
| HW self-efficacy | -.05 | (.16) | .770 | -.17 | (.13) | .183 | -.10 | (.13) | .465 | -.20 | (.10) | .051 | -.08 | (.12) | .539 | -.17 | (.10) | .098 |
| Intrinsic value | .04 | (.14) | .761 | .09 | (.24) | .708 | -.03 | (.11) | .811 | -.07 | (.20) | .718 | -.04 | (.10) | .711 | -.09 | (.19) | .630 |
| Utility value | .26 | (.14) | .066 | .32 | (.17) | .058 | .42 | (.11) | .000 | .52 | (.12) | .000 | .46 | (.11) | .000 | .54 | (.12) | .000 |
| Class' math valuing | -.06 | (.12) | .650 | .00 | (.22) | .984 | -.01 | (.10) | .892 | -.02 | (.13) | .885 | .00 | (.09) | .966 | .00 | (.12) | .977 |
| Disruptions in class | -.03 | (.12) | .823 | -.05 | (.14) | .739 | .08 | (.08) | .354 | .02 | (.11) | .842 | .07 | (.08) | .376 | .03 | (.11) | .809 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .28 | (.07) | .000 | .22 | (.06) | .000 | .28 | (.07) | .000 | .21 | (.07) | .002 | .28 | (.07) | .000 | .21 | (.07) | .003 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .11 | (.05) | .018 | .09 | (.05) | .084 | .09 | (.05) | .084 | .04 | (.05) | .390 | .09 | (.05) | .055 | .04 | (.05) | .389 |
| Cognitive ability | .01 | (.02) | .706 | .03 | (.03) | .237 | .01 | (.02) | .580 | .02 | (.03) | .561 | .02 | (.02) | .509 | .02 | (.03) | .604 |
| Conscientiousness | -.03 | (.03) | .292 | -.05 | (.03) | .088 | -.02 | (.03) | .418 | -.04 | (.03) | .212 | -.02 | (.03) | .423 | -.04 | (.03) | .173 |
| Test score | -.01 | (.03) | .768 | -.03 | (.04) | .489 | -.02 | (.03) | .587 | -.02 | (.04) | .523 | -.02 | (.03) | .530 | -.03 | (.04) | .496 |
| Self-concept | -.07 | (.04) | .105 | .05 | (.05) | .261 | -.08 | (.05) | .134 | .04 | (.05) | .399 | -.08 | (.05) | .120 | .04 | (.05) | .404 |
| HW self-efficacy | .02 | (.03) | .496 | -.04 | (.04) | .342 | .04 | (.04) | .369 | -.02 | (.05) | .626 | .03 | (.04) | .434 | -.03 | (.05) | .549 |
| Intrinsic value | .07 | (.03) | .049 | .04 | (.04) | .352 | .07 | (.04) | .051 | .05 | (.04) | .212 | .08 | (.04) | .038 | .06 | (.04) | .162 |
| Utility value | .69 | (.03) | .000 | .62 | (.03) | .000 | .69 | (.03) | .000 | .60 | (.04) | .000 | .69 | (.04) | .000 | .60 | (.04) | .000 |
| Class' math valuing | .04 | (.04) | .228 | .06 | (.03) | .069 | .03 | (.04) | .392 | .06 | (.03) | .086 | .03 | (.04) | .413 | .06 | (.03) | .099 |
| Disruptions in class | .00 | (.03) | .947 | .03 | (.03) | .333 | -.02 | (.03) | .425 | .03 | (.04) | .465 | -.02 | (.03) | .433 | .03 | (.04) | .456 |
| Wald-$X^2$-test | $X^2$ | (*df*) | *p* | $X^2$ | (*df*) | *p* | $X^2$ | (*df*) | *p* | $X^2$ | (*df*) | *p* | $X^2$ | (*df*) | *p* | $X^2$ | (*df*) | *p* |
| $\beta_{NC} = \beta_C$ | .08 | (1) | .782 | .29 | (1) | .591 | .06 | (1) | .801 | .71 | (1) | .398 | .14 | (1) | .711 | .65 | (1) | .419 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table E2
*Intervention Effects on Students' Math Utility Value Beliefs in the Quotations Condition Depending on Students' Responsiveness: Cutoff on Responsiveness Index above Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 8 | | | | | | 9 | | | | | | 10 | | | | | |
| C min. value | 9 | | | | | | 10 | | | | | | 11 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 462 (39%) | | | | | | 944 (79%) | | | | | | 1123 (94%) | | | | | |
| Complier | 734 (61%) | | | | | | 252 (21%) | | | | | | 73 (6%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .08 | (.10) | .446 | .32 | (.13) | .015 | .16 | (.08) | .031 | .20 | (.11) | .063 | .21 | (.06) | .000 | .17 | (.06) | .002 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .19 | (.08) | .022 | .09 | (.09) | .316 | .13 | (.05) | .020 | .08 | (.06) | .245 | .12 | (.04) | .003 | .07 | (.04) | .092 |
| Cognitive ability | -.02 | (.06) | .736 | .04 | (.06) | .584 | .01 | (.02) | .821 | .04 | (.03) | .150 | -.02 | (.02) | .346 | .00 | (.03) | .964 |
| Conscientiousness | .00 | (.05) | .969 | .00 | (.06) | .952 | -.01 | (.03) | .766 | .00 | (.04) | .974 | .00 | (.03) | .972 | -.01 | (.03) | .784 |
| Test score | .02 | (.04) | .578 | .03 | (.05) | .614 | .01 | (.03) | .757 | .00 | (.04) | .994 | .01 | (.03) | .637 | .00 | (.03) | .999 |
| Self-concept | -.06 | (.08) | .460 | .11 | (.10) | .254 | -.08 | (.05) | .089 | .10 | (.06) | .093 | -.07 | (.04) | .052 | .06 | (.05) | .216 |
| HW self-efficacy | .08 | (.08) | .288 | -.04 | (.07) | .552 | .06 | (.05) | .178 | -.06 | (.05) | .213 | .03 | (.03) | .340 | -.06 | (.03) | .062 |
| Intrinsic value | .08 | (.05) | .132 | -.06 | (.07) | .382 | .07 | (.04) | .058 | .01 | (.05) | .866 | .07 | (.03) | .015 | .05 | (.04) | .203 |
| Utility value | .62 | (.06) | .000 | .58 | (.06) | .000 | .66 | (.04) | .000 | .57 | (.04) | .000 | .66 | (.03) | .000 | .59 | (.03) | .000 |
| Class' math valuing | .00 | (.04) | .997 | .06 | (.06) | .338 | .03 | (.03) | .282 | .06 | (.04) | .129 | .03 | (.03) | .386 | .06 | (.04) | .081 |
| Disruptions in class | .06 | (.04) | .148 | .09 | (.05) | .096 | .03 | (.03) | .319 | .05 | (.05) | .313 | .01 | (.02) | .608 | .03 | (.04) | .487 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .45 | (.08) | .000 | .20 | (.09) | .026 | .56 | (.14) | .000 | .35 | (.16) | .032 | 1.33 | (.25) | .000 | 1.33 | (.37) | .000 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .03 | (.07) | .649 | .05 | (.06) | .471 | .04 | (.13) | .776 | .05 | (.09) | .615 | -.14 | (.19) | .472 | .06 | (.25) | .814 |
| Cognitive ability | -.04 | (.04) | .319 | -.03 | (.05) | .533 | -.10 | (.06) | .082 | -.14 | (.04) | .000 | .00 | (.11) | .977 | .14 | (.32) | .675 |
| Conscientiousness | -.01 | (.05) | .902 | -.06 | (.05) | .165 | .01 | (.06) | .937 | -.11 | (.05) | .033 | .16 | (.11) | .132 | -.01 | (.18) | .974 |
| Test score | .02 | (.04) | .658 | -.02 | (.05) | .616 | .02 | (.07) | .750 | .01 | (.09) | .916 | -.03 | (.12) | .793 | -.07 | (.13) | .581 |
| Self-concept | -.02 | (.09) | .848 | .06 | (.08) | .475 | .07 | (.13) | .604 | -.01 | (.12) | .912 | .26 | (.16) | .109 | -.15 | (.25) | .539 |
| HW self-efficacy | -.07 | (.05) | .184 | -.09 | (.05) | .076 | -.14 | (.08) | .099 | -.05 | (.09) | .593 | -.30 | (.16) | .056 | .00 | (.15) | .978 |
| Intrinsic value | .05 | (.05) | .315 | .08 | (.06) | .219 | .05 | (.07) | .538 | .06 | (.09) | .481 | .11 | (.20) | .571 | .20 | (.23) | .383 |
| Utility value | .67 | (.06) | .000 | .61 | (.06) | .000 | .61 | (.08) | .000 | .67 | (.06) | .000 | .44 | (.13) | .001 | .56 | (.12) | .000 |
| Class' math valuing | .09 | (.05) | .055 | .07 | (.04) | .112 | .06 | (.06) | .336 | .07 | (.07) | .286 | .01 | (.15) | .923 | .02 | (.12) | .874 |
| Disruptions in class | -.07 | (.04) | .120 | -.03 | (.04) | .532 | -.08 | (.06) | .182 | -.04 | (.05) | .492 | -.13 | (.08) | .097 | .06 | (.10) | .521 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | 7.13 | (1) | .008 | .37 | (1) | .542 | 5.13 | (1) | .024 | .38 | (1) | .538 | 23.04 | (1) | .000 | 10.03 | (1) | .002 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table E3
*Intervention Effects on Students' Math Utility Value Beliefs in the Text Condition Depending on Students' Responsiveness: Cutoff on Responsiveness Index below Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 4 | | | | | | 5 | | | | | | 6 | | | | | |
| C min. value | 5 | | | | | | 6 | | | | | | 7 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 77 (6%) | | | | | | 143 (11%) | | | | | | 217 (16%) | | | | | |
| Complier | 1278 (94%) | | | | | | 1212 (89%) | | | | | | 1138 (84%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Text | -.47 | (.34) | .165 | -.09 | (.53) | .873 | -.27 | (.16) | .096 | -.22 | (.23) | .341 | -.13 | (.17) | .439 | .10 | (.29) | .730 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.09 | (.18) | .635 | .14 | (.29) | .644 | -.03 | (.10) | .775 | .16 | (.12) | .181 | -.10 | (.11) | .387 | .03 | (.14) | .848 |
| Cognitive ability | -.17 | (.10) | .099 | -.14 | (.12) | .263 | -.11 | (.07) | .120 | -.16 | (.09) | .080 | -.09 | (.06) | .116 | -.15 | (.08) | .046 |
| Conscientiousness | -.14 | (.12) | .276 | -.29 | (.16) | .065 | -.07 | (.10) | .476 | -.17 | (.08) | .043 | -.04 | (.08) | .570 | -.10 | (.08) | .232 |
| Test score | -.08 | (.08) | .284 | -.08 | (.13) | .578 | -.05 | (.05) | .383 | -.04 | (.07) | .638 | -.01 | (.07) | .886 | .05 | (.08) | .539 |
| Self-concept | -.05 | (.14) | .753 | .21 | (.21) | .326 | .04 | (.10) | .711 | .12 | (.10) | .251 | .03 | (.11) | .780 | .08 | (.15) | .584 |
| HW self-efficacy | .21 | (.09) | .025 | .06 | (.15) | .700 | .09 | (.06) | .150 | .04 | (.06) | .525 | .08 | (.07) | .224 | .09 | (.12) | .430 |
| Intrinsic value | .25 | (.23) | .275 | -.25 | (.24) | .308 | .15 | (.10) | .109 | -.06 | (.11) | .572 | .04 | (.12) | .758 | -.10 | (.11) | .373 |
| Utility value | .56 | (.17) | .001 | .37 | (.17) | .035 | .59 | (.10) | .000 | .45 | (.13) | .001 | .66 | (.10) | .000 | .49 | (.10) | .000 |
| Class' math valuing | -.05 | (.11) | .617 | .31 | (.10) | .002 | -.02 | (.08) | .815 | .26 | (.09) | .003 | -.04 | (.07) | .596 | .17 | (.10) | .070 |
| Disruptions in class | -.01 | (.11) | .899 | -.02 | (.13) | .846 | .06 | (.08) | .415 | .04 | (.06) | .481 | .02 | (.07) | .744 | -.05 | (.07) | .480 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .18 | (.06) | .001 | .16 | (.07) | .027 | .20 | (.06) | .001 | .20 | (.07) | .003 | .20 | (.07) | .003 | .13 | (.10) | .169 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.03 | (.05) | .582 | .05 | (.05) | .296 | -.01 | (.06) | .827 | .03 | (.05) | .595 | .03 | (.07) | .710 | .09 | (.06) | .135 |
| Cognitive ability | -.03 | (.02) | .235 | -.03 | (.03) | .299 | -.03 | (.03) | .260 | -.02 | (.03) | .550 | -.02 | (.03) | .431 | .00 | (.03) | .903 |
| Conscientiousness | -.02 | (.03) | .380 | -.02 | (.03) | .440 | -.03 | (.03) | .287 | -.02 | (.03) | .400 | -.04 | (.03) | .206 | -.03 | (.03) | .361 |
| Test score | .06 | (.03) | .018 | .06 | (.03) | .074 | .07 | (.03) | .011 | .07 | (.03) | .053 | .06 | (.03) | .028 | .04 | (.03) | .168 |
| Self-concept | .00 | (.04) | .919 | -.01 | (.04) | .909 | -.01 | (.05) | .810 | -.02 | (.04) | .718 | -.01 | (.06) | .912 | -.01 | (.06) | .803 |
| HW self-efficacy | .00 | (.03) | .908 | .02 | (.03) | .514 | .01 | (.03) | .739 | .02 | (.03) | .443 | .00 | (.03) | .920 | .00 | (.04) | .930 |
| Intrinsic value | .02 | (.04) | .574 | .03 | (.03) | .314 | .02 | (.05) | .738 | .04 | (.03) | .248 | .02 | (.05) | .671 | .04 | (.04) | .305 |
| Utility value | .66 | (.03) | .000 | .59 | (.04) | .000 | .66 | (.03) | .000 | .59 | (.04) | .000 | .64 | (.04) | .000 | .58 | (.05) | .000 |
| Class' math valuing | .05 | (.03) | .164 | .08 | (.03) | .017 | .05 | (.04) | .205 | .07 | (.03) | .039 | .05 | (.04) | .175 | .07 | (.04) | .059 |
| Disruptions in class | -.04 | (.03) | .165 | -.04 | (.03) | .202 | -.05 | (.03) | .106 | -.05 | (.03) | .112 | -.05 | (.03) | .134 | -.02 | (.03) | .481 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | 3.38 | (1) | .066 | .18 | (1) | .675 | 6.79 | (1) | .009 | 2.57 | (1) | .109 | 2.79 | (1) | .095 | .01 | (1) | .923 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table E4
*Intervention Effects on Students' Math Utility Value Beliefs in the Text Condition Depending on Students' Responsiveness: Cutoff on Responsiveness Index above Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 8 | | | | | | 9 | | | | | | 10 | | | | | |
| C min. value | 9 | | | | | | 10 | | | | | | 11 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 312 (23%) | | | | | | 713 (53%) | | | | | | 1208 (89%) | | | | | |
| Complier | 1043 (77%) | | | | | | 642 (47%) | | | | | | 147 (11%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Text | -.04 | (.15) | .800 | .14 | (.18) | .441 | .06 | (.11) | .589 | .17 | (.12) | .180 | .10 | (.07) | .168 | .15 | (.09) | .092 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.11 | (.10) | .254 | .07 | (.11) | .522 | -.15 | (.09) | .109 | .04 | (.09) | .664 | -.04 | (.06) | .492 | .07 | (.06) | .223 |
| Cognitive ability | -.10 | (.05) | .051 | -.14 | (.07) | .050 | -.05 | (.03) | .124 | -.08 | (.04) | .042 | -.03 | (.02) | .137 | -.04 | (.03) | .180 |
| Conscientiousness | -.08 | (.07) | .258 | -.11 | (.06) | .069 | -.07 | (.05) | .189 | -.12 | (.04) | .003 | -.04 | (.04) | .332 | -.06 | (.03) | .053 |
| Test score | .03 | (.05) | .611 | .03 | (.08) | .662 | .01 | (.04) | .845 | .01 | (.05) | .831 | .05 | (.03) | .107 | .03 | (.04) | .553 |
| Self-concept | .03 | (.11) | .787 | .08 | (.11) | .477 | .02 | (.08) | .811 | .07 | (.08) | .440 | -.02 | (.05) | .678 | .02 | (.06) | .804 |
| HW self-efficacy | .07 | (.06) | .239 | .09 | (.08) | .261 | .04 | (.05) | .429 | .07 | (.06) | .215 | .01 | (.03) | .820 | .01 | (.04) | .777 |
| Intrinsic value | .02 | (.09) | .799 | -.07 | (.10) | .468 | .04 | (.06) | .553 | -.02 | (.07) | .817 | .05 | (.05) | .249 | .01 | (.05) | .811 |
| Utility value | .65 | (.08) | .000 | .47 | (.09) | .000 | .65 | (.05) | .000 | .54 | (.06) | .000 | .62 | (.04) | .000 | .55 | (.04) | .000 |
| Class' math valuing | .01 | (.06) | .859 | .16 | (.08) | .043 | .02 | (.05) | .755 | .12 | (.05) | .012 | .05 | (.05) | .284 | .10 | (.04) | .003 |
| Disruptions in class | .01 | (.07) | .922 | -.02 | (.07) | .750 | -.01 | (.04) | .792 | -.05 | (.05) | .284 | -.02 | (.04) | .596 | -.04 | (.04) | .226 |
| Complier | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .19 | (.08) | .016 | .13 | (.09) | .141 | .19 | (.11) | .101 | .09 | (.12) | .445 | .22 | (.17) | .179 | .07 | (.24) | .783 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .05 | (.08) | .503 | .06 | (.06) | .298 | .14 | (.11) | .214 | .08 | (.08) | .300 | .11 | (.12) | .344 | -.02 | (.10) | .884 |
| Cognitive ability | -.01 | (.03) | .753 | .01 | (.03) | .739 | -.03 | (.04) | .481 | .01 | (.04) | .833 | -.10 | (.07) | .188 | -.05 | (.08) | .565 |
| Conscientiousness | -.02 | (.03) | .490 | -.02 | (.03) | .526 | -.02 | (.05) | .614 | .02 | (.04) | .647 | -.05 | (.11) | .655 | -.01 | (.10) | .941 |
| Test score | .05 | (.03) | .090 | .04 | (.03) | .153 | .09 | (.04) | .024 | .07 | (.03) | .020 | .05 | (.06) | .412 | .12 | (.07) | .067 |
| Self-concept | -.01 | (.06) | .916 | -.01 | (.06) | .805 | -.01 | (.09) | .906 | -.04 | (.08) | .652 | .11 | (.14) | .421 | .02 | (.15) | .896 |
| HW self-efficacy | -.01 | (.04) | .799 | -.03 | (.04) | .467 | .00 | (.05) | .940 | -.04 | (.05) | .441 | .06 | (.07) | .433 | .07 | (.08) | .405 |
| Intrinsic value | .03 | (.05) | .575 | .06 | (.05) | .183 | .02 | (.07) | .783 | .05 | (.07) | .435 | -.05 | (.08) | .548 | .02 | (.11) | .879 |
| Utility value | .64 | (.04) | .000 | .60 | (.04) | .000 | .62 | (.06) | .000 | .57 | (.06) | .000 | .70 | (.09) | .000 | .62 | (.07) | .000 |
| Class' math valuing | .04 | (.05) | .420 | .07 | (.04) | .111 | .06 | (.07) | .408 | .08 | (.05) | .129 | .02 | (.08) | .851 | .11 | (.09) | .227 |
| Disruptions in class | -.05 | (.03) | .169 | -.03 | (.03) | .289 | -.05 | (.04) | .237 | -.02 | (.04) | .632 | -.06 | (.05) | .212 | .02 | (.05) | .764 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_{C}$ | 1.35 | (1) | .245 | .00 | (1) | .966 | .40 | (1) | .527 | .13 | (1) | .720 | .35 | (1) | .556 | .08 | (1) | .782 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table E5
*Intervention Effects on Students' Math Utility Value Beliefs in the Quotations Condition Depending on Students' Responsiveness: Cutoff on Extended Responsiveness Index below Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 4 | | | | | | 5 | | | | | | 6 | | | | | |
| C min. value | 5 | | | | | | 6 | | | | | | 7 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 73 (6%) | | | | | | 106 (9%) | | | | | | 137 (11%) | | | | | |
| Complier | 1123 (94%) | | | | | | 1090 (87%) | | | | | | 1059 (89%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .24 | (.36) | .499 | .19 | (.45) | .678 | .28 | (.27) | .295 | .39 | (.29) | .179 | .38 | (.24) | .118 | .42 | (.25) | .093 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .05 | (.35) | .893 | -.02 | (.31) | .961 | .14 | (.27) | .595 | .10 | (.24) | .670 | .09 | (.23) | .687 | .16 | (.19) | .407 |
| Cognitive ability | -.48 | (.25) | .056 | -.34 | (.24) | .158 | -.32 | (.11) | .003 | -.17 | (.13) | .195 | -.24 | (.08) | .003 | -.09 | (.09) | .302 |
| Conscientiousness | .15 | (.15) | .326 | .03 | (.20) | .868 | .12 | (.10) | .234 | .06 | (.11) | .617 | .06 | (.09) | .498 | .02 | (.09) | .873 |
| Test score | .21 | (.32) | .511 | .25 | (.31) | .410 | .16 | (.13) | .211 | .15 | (.18) | .398 | .08 | (.11) | .501 | .04 | (.15) | .813 |
| Self-concept | .14 | (.19) | .453 | .07 | (.25) | .770 | .18 | (.13) | .170 | .16 | (.17) | .328 | .16 | (.13) | .222 | .18 | (.15) | .232 |
| HW self-efficacy | -.01 | (.23) | .980 | -.16 | (.15) | .305 | -.06 | (.14) | .686 | -.16 | (.12) | .169 | -.09 | (.13) | .500 | -.20 | (.10) | .049 |
| Intrinsic value | .04 | (.25) | .865 | .13 | (.31) | .674 | -.08 | (.13) | .576 | -.10 | (.23) | .659 | -.03 | (.11) | .767 | -.07 | (.20) | .730 |
| Utility value | .18 | (.16) | .235 | .21 | (.19) | .272 | .37 | (.12) | .003 | .43 | (.15) | .004 | .45 | (.11) | .000 | .51 | (.12) | .000 |
| Class' math valuing | .03 | (.14) | .815 | .15 | (.23) | .510 | .06 | (.14) | .679 | .10 | (.19) | .587 | -.01 | (.10) | .940 | -.01 | (.13) | .960 |
| Disruptions in class | .01 | (.18) | .949 | .05 | (.19) | .801 | .08 | (.09) | .383 | .04 | (.14) | .794 | .09 | (.08) | .302 | .03 | (.12) | .813 |
| Complier | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .30 | (.07) | .000 | .24 | (.06) | .000 | .29 | (.07) | .000 | .22 | (.07) | .001 | .28 | (.07) | .000 | .22 | (.07) | .001 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .10 | (.05) | .052 | .07 | (.05) | .162 | .09 | (.05) | .076 | .06 | (.05) | .258 | .10 | (.05) | .056 | .05 | (.05) | .376 |
| Cognitive ability | .00 | (.02) | .976 | .02 | (.02) | .293 | .01 | (.02) | .561 | .03 | (.03) | .332 | .01 | (.02) | .559 | .02 | (.03) | .572 |
| Conscientiousness | -.03 | (.03) | .405 | -.04 | (.03) | .151 | -.03 | (.03) | .339 | -.05 | (.03) | .128 | -.03 | (.03) | .411 | -.04 | (.03) | .187 |
| Test score | -.01 | (.03) | .754 | -.03 | (.04) | .491 | -.02 | (.03) | .517 | -.03 | (.04) | .455 | -.02 | (.03) | .529 | -.02 | (.04) | .542 |
| Self-concept | -.07 | (.04) | .119 | .05 | (.05) | .276 | -.08 | (.05) | .082 | .04 | (.05) | .440 | -.08 | (.05) | .133 | .04 | (.05) | .447 |
| HW self-efficacy | .01 | (.03) | .695 | -.04 | (.04) | .262 | .03 | (.04) | .433 | -.03 | (.04) | .457 | .04 | (.04) | .371 | -.02 | (.05) | .664 |
| Intrinsic value | .07 | (.03) | .035 | .04 | (.04) | .329 | .08 | (.04) | .023 | .06 | (.04) | .136 | .07 | (.04) | .051 | .05 | (.04) | .212 |
| Utility value | .69 | (.03) | .000 | .63 | (.03) | .000 | .69 | (.03) | .000 | .62 | (.03) | .000 | .69 | (.04) | .000 | .61 | (.04) | .000 |
| Class' math valuing | .04 | (.03) | .267 | .05 | (.03) | .090 | .03 | (.04) | .437 | .04 | (.03) | .161 | .03 | (.04) | .418 | .06 | (.03) | .099 |
| Disruptions in class | -.01 | (.03) | .805 | .02 | (.03) | .579 | -.02 | (.03) | .507 | .02 | (.04) | .547 | -.03 | (.03) | .398 | .02 | (.03) | .480 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_{C}$ | .02 | (1) | .885 | .02 | (1) | .901 | .00 | (1) | .962 | .27 | (1) | .601 | .13 | (1) | .720 | .48 | (1) | .488 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

## Part F) Information on the intervention boosters to be completed at home

### Background information

Students' booster tasks completed at home after the in-class intervention were collected and analyzed ($N_{total}$ = 1311). Each booster consisted of a preformatted A4 page including two to three questions with the answers requiring only one word or a few short sentences. Because of their brevity, it was impossible to apply the same theory-driven coding procedure to the boosters as used for the relevance essays. The boosters were thus only coded on whether they were completed as intended or not. Information on the coding of these boosters, the frequency distributions, the intercorrelations with other variables under investigation, and the predictors can be found in Tables F1 to F3 and Figure F1. Information on the CACE analyses using an extended responsiveness index including data on the boosters can be found in Tables F4 to F9 and Figures F2 and B3.

### Task of intervention booster 1 (both conditions)

1) Please recall the math lesson during which you worked on an assignment about the importance of math. What was your task?
2) Which arguments can you remember?
3) Can you remember which argument you found most convincing? (…) Please write down the argument which is most important for you.

### Tasks of intervention booster 2

**Text condition:**

1) Please think of a person you know personally and to whom math is useful, e.g., for his/her job or leisure activities. Who do you think of? Please just write down your relation to this person, e.g. my aunt.
2) What is this person's profession?
3) Why is or was math important in this person's life? Please note down two reasons.

**Quotations condition:**

1) Please go to a computer and browse the webpage http://du-kannst-mathe.de/. Which argument on the webpage about the utility of mathematics convinces you the most?
2) Explain why you (…) find this argument most convincing for you personally?

### Coding procedure of the intervention boosters (both conditions)

The intervention boosters were coded on whether they were completed as intended or not, resulting in a dummy variable comprising the values 0 (booster not turned in or not completed as intended) and 1 (booster completed as intended) for each of the two boosters. A scale representing the number of boosters completed as intended (0, 1, or 2) was calculated by summing the two dummy variables. The descriptive statistics of the intervention boosters are presented below.

### Descriptives on the intervention boosters per intervention condition

Table F1

*Number, means, and standard deviations of the intervention boosters*

| | *N* | *M* | *SD* |
|---|---|---|---|
| Quotations | 561 | 0.76 | 0.83 |
| Text | 718 | 0.89 | 0.87 |

*Notes.* *N* = number; *M* = mean; *SD* = standard deviation.

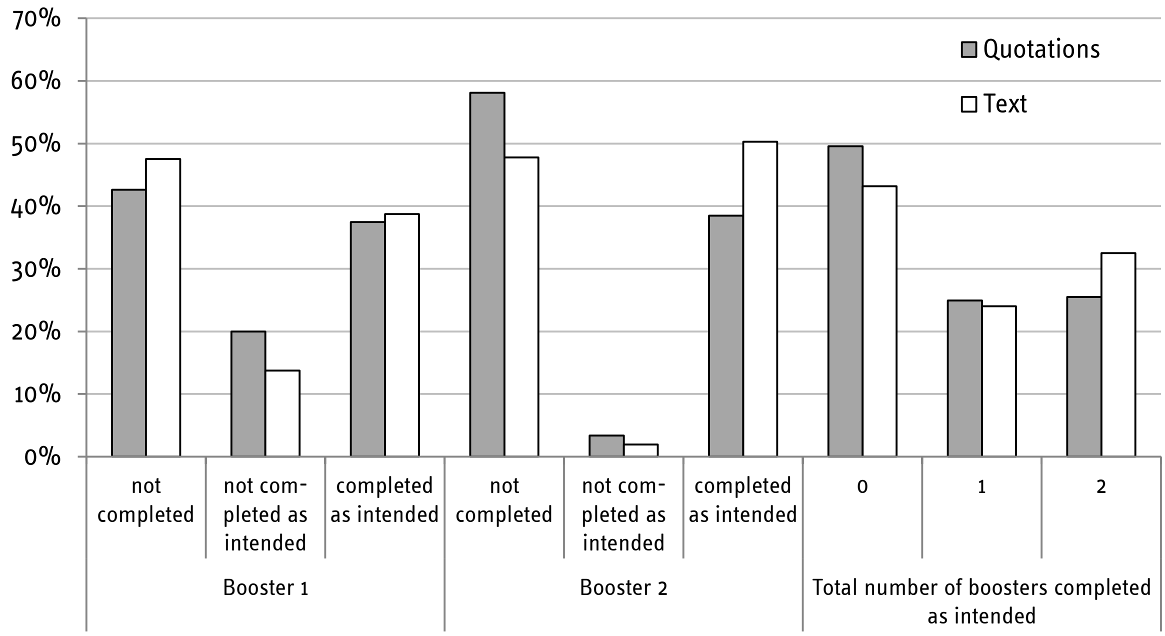



**Figure F1:** Frequency distributions: students' completion of the intervention boosters per intervention condition.

## Intercorrelations (including the intervention boosters)

Table F2

*Correlations between students' completion of the boosters, the responsiveness index, students' individual characteristics, and classroom perceptions at T1, as well as math-related utility value at T2 and T3 in the quotations condition (below diagonal) and in the text condition (above diagonal)*

| | | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | | (7) | | (8) | | (9) | | (10) | | (11) | | (12) | | (13) | | (14) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* | *r* | *p* |
| (1) | Completed Boosters | - | - | .18 | .000 | -.27 | .000 | .11 | .005 | .26 | .000 | .26 | .000 | .16 | .000 | .08 | .035 | .13 | .003 | .15 | .007 | .04 | .414 | -.13 | .007 | .20 | .000 | .19 | .000 |
| (2) | Responsiveness index | .15 | .001 | - | - | -.12 | .001 | .10 | .021 | .20 | .000 | .17 | .001 | .18 | .000 | .18 | .000 | .23 | .000 | .25 | .000 | .17 | .001 | -.13 | .020 | .31 | .000 | .24 | .000 |
| (3) | Gender (1 = male) | -.22 | .000 | -.03 | .473 | - | - | -.03 | .461 | -.17 | .000 | .02 | .662 | .17 | .000 | .09 | .024 | .16 | .000 | .05 | .216 | .13 | .001 | -.07 | .083 | .01 | .743 | .08 | .036 |
| (4) | Cognitive ability | .04 | .335 | .04 | .139 | -.01 | .797 | - | - | .01 | .761 | .30 | .000 | .36 | .000 | .12 | .007 | .25 | .000 | .11 | .005 | .02 | .720 | -.05 | .348 | .07 | .076 | .02 | .698 |
| (5) | Conscientiousness | .27 | .000 | .17 | .001 | -.09 | .034 | -.11 | .013 | - | - | .16 | .000 | .21 | .000 | .22 | .000 | .20 | .000 | .25 | .000 | .10 | .025 | -.13 | .000 | .18 | .000 | .16 | .000 |
| (6) | Test score | .19 | .000 | .20 | .000 | .12 | .006 | .25 | .000 | .10 | .054 | - | - | .51 | .000 | .27 | .000 | .42 | .000 | .18 | .000 | .04 | .518 | -.09 | .175 | .20 | .000 | .18 | .000 |
| (7) | Self-concept | .06 | .238 | .12 | .039 | .28 | .000 | .18 | .000 | .20 | .001 | .49 | .000 | - | - | .47 | .000 | .74 | .000 | .38 | .000 | .16 | .000 | -.11 | .003 | .31 | .000 | .23 | .000 |
| (8) | HW self-efficacy | .09 | .155 | .15 | .001 | .09 | .051 | .02 | .686 | .33 | .000 | .24 | .000 | .47 | .000 | - | - | .44 | .000 | .27 | .000 | .20 | .000 | -.13 | .016 | .26 | .000 | .24 | .000 |
| (9) | Intrinsic value | .10 | .056 | .21 | .000 | .14 | .002 | .12 | .003 | .29 | .000 | .43 | .000 | .71 | .000 | .45 | .000 | - | - | .50 | .000 | .36 | .000 | -.17 | .011 | .40 | .000 | .33 | .000 |
| (10) | Utility value | .11 | .027 | .20 | .001 | .12 | .006 | -.02 | .593 | .33 | .000 | .26 | .000 | .41 | .000 | .37 | .000 | .55 | .000 | - | - | .36 | .000 | -.17 | .009 | .65 | .000 | .56 | .000 |
| (11) | Class' math valuing | .04 | .349 | .14 | .005 | -.07 | .107 | -.01 | .837 | .09 | .120 | .11 | .056 | .17 | .002 | .22 | .000 | .30 | .000 | .29 | .000 | - | - | -.40 | .000 | .28 | .000 | .31 | .000 |
| (12) | Disruptions in class | -.07 | .113 | -.07 | .185 | .04 | .337 | .04 | .467 | -.09 | .109 | -.06 | .162 | -.04 | .368 | -.06 | .215 | -.09 | .038 | -.12 | .005 | -.23 | .012 | - | - | -.19 | .003 | -.20 | .004 |
| (13) | Utility value T2 | .14 | .001 | .19 | .000 | .11 | .009 | -.05 | .197 | .25 | .000 | .17 | .000 | .29 | .000 | .29 | .000 | .42 | .000 | .70 | .000 | .22 | .000 | -.09 | .008 | - | - | .64 | .000 |
| (14) | Utility value T3 | .09 | .003 | .16 | .001 | .07 | .111 | -.01 | .761 | .20 | .000 | .17 | .000 | .27 | .000 | .18 | .000 | .34 | .000 | .59 | .000 | .16 | .000 | .00 | .984 | .67 | .000 | - | - |

*Notes*. T = time; HW = homework; *r* = Pearson's correlation coefficient; *p* = *p*-value.

**Predicting students' completion of the intervention boosters through students' individual characteristics and classroom perceptions per intervention condition**

Table F3

*Predicting the number of completed intervention boosters through students' individual characteristics and classroom perceptions*

| | Quotations | | | Text | | |
|---|---|---|---|---|---|---|
| | *β* | *(SE)* | *p* | *β* | *(SE)* | *p* |
| Basic characteristics | | | | | | |
| Gender (1 = male) | -.43 | (.08) | .000 | -.54 | (.08) | .000 |
| Cognitive ability | .02 | (.04) | .652 | .02 | (.04) | .507 |
| Conscientiousness | .24 | (.05) | .000 | .16 | (.05) | .001 |
| Math achievement | | | | | | |
| Test score | .23 | (.06) | .000 | .17 | (.06) | .003 |
| Math motivation | | | | | | |
| Self-concept | -.04 | (.07) | .584 | .07 | (.07) | .320 |
| HW self-efficacy | .00 | (.07) | .964 | -.03 | (.05) | .478 |
| Intrinsic value | -.03 | (.06) | .619 | -.01 | (.07) | .842 |
| Utility value | .04 | (.05) | .365 | .06 | (.05) | .225 |
| Classroom perceptions | | | | | | |
| Class' math valuing | -.02 | (.05) | .669 | .00 | (.05) | .943 |
| Disruptions in class | -.04 | (.04) | .334 | -.10 | (.03) | .002 |

*Notes.* HW = homework; *β* = standardized regression coefficient; *SE* = standard error; *p* = *p*-value.

**Creation of the extended responsiveness index (including data on the boosters)**

To check if the results of the CACE-analyses were stable, we added the number of completed intervention assignments to the original responsiveness index, resulting in the creation of an extended responsiveness index comprising values from 1 to 13. We used the extended responsiveness index to conduct robustness CACE-analyses.

Table F4

*Creation of the extended responsiveness index*

| Value on extended index | Value on responsiveness index + Number of boosters completed as intended |
|---|---|
| 1 | 1+0 |
| 2 | 2+0; 1+1 |
| 3 | 3+0; 2+1; 1+2 |
| 4 | 4+0; 3+1; 2+2 |
| 5 | 5+0; 4+1; 3+2 |
| 6 | 6+0; 5+1; 4+2 |
| 7 | 7+0; 6+1; 5+2 |
| 8 | 8+0; 7+1; 6+2 |
| 9 | 9+0; 8+1; 7+2 |
| 10 | 10+0; 9+1; 8+2 |
| 11 | 11+0; 10+1; 9+2 |
| 12 | 11+1; 10+2 |
| 13 | 11+2 |

## Results of the CACE analyses using the extended responsiveness index

Table F5

*Intervention Effects on Students' Math Utility Value Beliefs in the Quotations Condition Depending on Students' Responsiveness: Cutoff on Extended Responsiveness Index below Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 4 | | | | | | 5 | | | | | | 6 | | | | | |
| C min. value | 5 | | | | | | 6 | | | | | | 7 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 73 (6%) | | | | | | 106 (9%) | | | | | | 137 (11%) | | | | | |
| Complier | 1123 (94%) | | | | | | 1090 (87%) | | | | | | 1059 (89%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .24 | (.36) | .499 | .19 | (.45) | .678 | .28 | (.27) | .295 | .39 | (.29) | .179 | .38 | (.24) | .118 | .42 | (.25) | .093 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .05 | (.35) | .893 | -.02 | (.31) | .961 | .14 | (.27) | .595 | .10 | (.24) | .670 | .09 | (.23) | .687 | .16 | (.19) | .407 |
| Cognitive ability | -.48 | (.25) | .056 | -.34 | (.24) | .158 | -.32 | (.11) | .003 | -.17 | (.13) | .195 | -.24 | (.08) | .003 | -.09 | (.09) | .302 |
| Conscientiousness | .15 | (.15) | .326 | .03 | (.20) | .868 | .12 | (.10) | .234 | .06 | (.11) | .617 | .06 | (.09) | .498 | .02 | (.09) | .873 |
| Test score | .21 | (.32) | .511 | .25 | (.31) | .410 | .16 | (.13) | .211 | .15 | (.18) | .398 | .08 | (.11) | .501 | .04 | (.15) | .813 |
| Self-concept | .14 | (.19) | .453 | .07 | (.25) | .770 | .18 | (.13) | .170 | .16 | (.17) | .328 | .16 | (.13) | .222 | .18 | (.15) | .232 |
| HW self-efficacy | -.01 | (.23) | .980 | -.16 | (.15) | .305 | -.06 | (.14) | .686 | -.16 | (.12) | .169 | -.09 | (.13) | .500 | -.20 | (.10) | .049 |
| Intrinsic value | .04 | (.25) | .865 | .13 | (.31) | .674 | -.08 | (.13) | .576 | -.10 | (.23) | .659 | -.03 | (.11) | .767 | -.07 | (.20) | .730 |
| Utility value | .18 | (.16) | .235 | .21 | (.19) | .272 | .37 | (.12) | .003 | .43 | (.15) | .004 | .45 | (.11) | .000 | .51 | (.12) | .000 |
| Class' math valuing | .03 | (.14) | .815 | .15 | (.23) | .510 | .06 | (.14) | .679 | .10 | (.19) | .587 | -.01 | (.10) | .940 | -.01 | (.13) | .960 |
| Disruptions in class | .01 | (.18) | .949 | .05 | (.19) | .801 | .08 | (.09) | .383 | .04 | (.14) | .794 | .09 | (.08) | .302 | .03 | (.12) | .813 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .30 | (.07) | .000 | .24 | (.06) | .000 | .29 | (.07) | .000 | .22 | (.07) | .001 | .28 | (.07) | .000 | .22 | (.07) | .001 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .10 | (.05) | .052 | .07 | (.05) | .162 | .09 | (.05) | .076 | .06 | (.05) | .258 | .10 | (.05) | .056 | .05 | (.05) | .376 |
| Cognitive ability | .00 | (.02) | .976 | .02 | (.02) | .293 | .01 | (.02) | .561 | .03 | (.03) | .332 | .01 | (.02) | .559 | .02 | (.03) | .572 |
| Conscientiousness | -.03 | (.03) | .405 | -.04 | (.03) | .151 | -.03 | (.03) | .339 | -.05 | (.03) | .128 | -.03 | (.03) | .411 | -.04 | (.03) | .187 |
| Test score | -.01 | (.03) | .754 | -.03 | (.04) | .491 | -.02 | (.03) | .517 | -.03 | (.04) | .455 | -.02 | (.03) | .529 | -.02 | (.04) | .542 |
| Self-concept | -.07 | (.04) | .119 | .05 | (.05) | .276 | -.08 | (.05) | .082 | .04 | (.05) | .440 | -.08 | (.05) | .133 | .04 | (.05) | .447 |
| HW self-efficacy | .01 | (.03) | .695 | -.04 | (.04) | .262 | .03 | (.04) | .433 | -.03 | (.04) | .457 | .04 | (.04) | .371 | -.02 | (.05) | .664 |
| Intrinsic value | .07 | (.03) | .035 | .04 | (.04) | .329 | .08 | (.04) | .023 | .06 | (.04) | .136 | .07 | (.04) | .051 | .05 | (.04) | .212 |
| Utility value | .69 | (.03) | .000 | .63 | (.03) | .000 | .69 | (.03) | .000 | .62 | (.03) | .000 | .69 | (.04) | .000 | .61 | (.04) | .000 |
| Class' math valuing | .04 | (.03) | .267 | .05 | (.03) | .090 | .03 | (.04) | .437 | .04 | (.03) | .161 | .03 | (.04) | .418 | .06 | (.03) | .099 |
| Disruptions in class | -.01 | (.03) | .805 | .02 | (.03) | .579 | -.02 | (.03) | .507 | .02 | (.04) | .547 | -.03 | (.03) | .398 | .02 | (.03) | .480 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | .02 | (1) | .885 | .02 | (1) | .901 | .00 | (1) | .962 | .27 | (1) | .601 | .13 | (1) | .720 | .48 | (1) | .488 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table F6

*Intervention Effects on Students' Math Utility Value Beliefs in the Quotations Condition Depending on Students' Responsiveness: Cutoff on Extended Responsiveness Index above Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 8 | | | | | | 9 | | | | | | 10 | | | | | |
| C min. value | 9 | | | | | | 10 | | | | | | 11 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 306 (26%) | | | | | | 615 (51%) | | | | | | 948 (79%) | | | | | |
| Complier | 890 (74%) | | | | | | 581 (49%) | | | | | | 248 (21%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .33 | (.15) | .029 | .34 | (.16) | .035 | .11 | (.10) | .232 | .24 | (.12) | .050 | .30 | (.09) | .000 | .32 | (.10) | .001 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .14 | (.13) | .256 | .13 | (.11) | .254 | .10 | (.08) | .228 | .07 | (.08) | .353 | .06 | (.07) | .341 | -.01 | (.05) | .829 |
| Cognitive ability | -.15 | (.05) | .001 | -.06 | (.05) | .276 | -.01 | (.03) | .740 | .01 | (.04) | .753 | -.07 | (.03) | .025 | -.02 | (.03) | .520 |
| Conscientiousness | .05 | (.06) | .369 | .01 | (.06) | .919 | -.01 | (.04) | .810 | -.02 | (.04) | .619 | .01 | (.04) | .908 | -.05 | (.04) | .172 |
| Test score | .03 | (.06) | .626 | .01 | (.07) | .894 | .01 | (.03) | .745 | .01 | (.04) | .756 | .01 | (.04) | .774 | -.01 | (.04) | .773 |
| Self-concept | .14 | (.08) | .062 | .20 | (.09) | .017 | -.05 | (.07) | .451 | .12 | (.07) | .081 | .06 | (.06) | .321 | .12 | (.06) | .052 |
| HW self-efficacy | -.06 | (.09) | .467 | -.13 | (.06) | .044 | .06 | (.07) | .387 | -.07 | (.06) | .220 | -.03 | (.05) | .533 | -.06 | (.05) | .180 |
| Intrinsic value | .01 | (.06) | .896 | -.12 | (.09) | .192 | .08 | (.05) | .095 | -.04 | (.06) | .530 | .03 | (.04) | .502 | .00 | (.05) | .931 |
| Utility value | .52 | (.07) | .000 | .53 | (.07) | .000 | .62 | (.05) | .000 | .54 | (.05) | .000 | .60 | (.04) | .000 | .56 | (.04) | .000 |
| Class' math valuing | .01 | (.04) | .724 | .06 | (.07) | .418 | .03 | (.03) | .294 | .08 | (.06) | .166 | .05 | (.03) | .143 | .07 | (.04) | .085 |
| Disruptions in class | .04 | (.05) | .484 | .04 | (.07) | .517 | .07 | (.04) | .097 | .05 | (.05) | .312 | .02 | (.04) | .636 | .04 | (.05) | .394 |
| Complier | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* | $\beta_{C}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .28 | (.08) | .000 | .21 | (.08) | .008 | .47 | (.11) | .000 | .27 | (.11) | .010 | .28 | (.16) | .088 | .08 | (.21) | .688 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .08 | (.06) | .194 | .04 | (.06) | .525 | .10 | (.10) | .334 | .06 | (.06) | .389 | .19 | (.11) | .095 | .23 | (.09) | .010 |
| Cognitive ability | .03 | (.03) | .366 | .03 | (.04) | .499 | -.06 | (.04) | .178 | -.05 | (.05) | .353 | .06 | (.05) | .247 | .05 | (.10) | .596 |
| Conscientiousness | -.04 | (.03) | .208 | -.06 | (.04) | .145 | -.01 | (.05) | .911 | -.06 | (.05) | .186 | -.08 | (.06) | .206 | .00 | (.09) | .987 |
| Test score | -.01 | (.03) | .821 | -.03 | (.04) | .459 | .02 | (.05) | .645 | -.03 | (.05) | .587 | -.01 | (.05) | .841 | -.03 | (.07) | .733 |
| Self-concept | -.14 | (.06) | .019 | -.01 | (.05) | .890 | .00 | (.11) | .997 | .03 | (.07) | .655 | -.28 | (.12) | .024 | -.05 | (.12) | .711 |
| HW self-efficacy | .05 | (.05) | .281 | -.02 | (.05) | .771 | -.06 | (.07) | .367 | -.04 | (.06) | .507 | .09 | (.08) | .258 | -.05 | (.14) | .715 |
| Intrinsic value | .08 | (.04) | .069 | .10 | (.05) | .026 | .02 | (.05) | .704 | .08 | (.06) | .187 | .12 | (.07) | .080 | .09 | (.09) | .335 |
| Utility value | .72 | (.04) | .000 | .63 | (.05) | .000 | .67 | (.07) | .000 | .64 | (.06) | .000 | .73 | (.07) | .000 | .62 | (.08) | .000 |
| Class' math valuing | .03 | (.04) | .517 | .03 | (.04) | .381 | .07 | (.05) | .220 | .05 | (.05) | .358 | -.01 | (.06) | .933 | .03 | (.06) | .623 |
| Disruptions in class | -.02 | (.03) | .459 | .01 | (.04) | .743 | -.08 | (.05) | .083 | -.01 | (.04) | .849 | -.07 | (.06) | .252 | -.05 | (.10) | .636 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_{C}$ | .05 | (1) | .826 | .38 | (1) | .537 | 4.56 | (1) | .033 | .03 | (1) | .864 | .01 | (1) | .904 | .75 | (1) | .386 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

Table F7

*Intervention Effects on Students' Math Utility Value Beliefs in the Text Condition Depending on Students' Responsiveness: Cutoff on Extended Responsiveness Index below Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 4 | | | | | | 5 | | | | | | 6 | | | | | |
| C min. value | 5 | | | | | | 6 | | | | | | 7 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 60 (4%) | | | | | | 111 (8%) | | | | | | 164 (12%) | | | | | |
| Complier | 1295 (96%) | | | | | | 1244 (92%) | | | | | | 1191 (88%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Text | -.77 | (.32) | .015 | -.42 | (.27) | .113 | -.43 | (.18) | .014 | -.38 | (.16) | .019 | -.25 | (.18) | .182 | -.13 | (.29) | .663 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .13 | (.20) | .533 | .26 | (.26) | .317 | .05 | (.13) | .692 | .24 | (.12) | .051 | -.08 | (.12) | .536 | .02 | (.20) | .920 |
| Cognitive ability | -.17 | (.09) | .063 | -.17 | (.09) | .051 | -.11 | (.11) | .315 | -.18 | (.09) | .030 | -.10 | (.07) | .144 | -.11 | (.06) | .037 |
| Conscientiousness | -.17 | (.10) | .088 | -.35 | (.10) | .001 | -.10 | (.09) | .264 | -.23 | (.08) | .004 | -.02 | (.08) | .787 | -.14 | (.08) | .072 |
| Test score | -.05 | (.08) | .552 | .14 | (.15) | .349 | .00 | (.06) | .992 | -.06 | (.09) | .557 | -.06 | (.07) | .413 | -.06 | (.09) | .526 |
| Self-concept | -.08 | (.12) | .534 | .19 | (.19) | .323 | -.02 | (.11) | .842 | .17 | (.12) | .142 | .01 | (.10) | .942 | .17 | (.12) | .172 |
| HW self-efficacy | .20 | (.08) | .011 | .05 | (.13) | .714 | .13 | (.07) | .085 | .03 | (.10) | .812 | .05 | (.06) | .389 | -.03 | (.09) | .778 |
| Intrinsic value | .10 | (.22) | .655 | -.42 | (.21) | .044 | .15 | (.14) | .295 | -.11 | (.14) | .414 | .02 | (.12) | .887 | -.17 | (.13) | .205 |
| Utility value | .71 | (.15) | .000 | .41 | (.13) | .002 | .65 | (.12) | .000 | .34 | (.11) | .003 | .68 | (.11) | .000 | .50 | (.12) | .000 |
| Class' math valuing | .05 | (.12) | .669 | .35 | (.08) | .000 | -.01 | (.09) | .892 | .29 | (.08) | .000 | -.02 | (.08) | .836 | .29 | (.10) | .004 |
| Disruptions in class | .02 | (.14) | .898 | -.05 | (.13) | .683 | .09 | (.09) | .339 | .00 | (.10) | .997 | .12 | (.08) | .134 | .00 | (.09) | .963 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .19 | (.05) | .000 | .18 | (.06) | .002 | .20 | (.06) | .001 | .20 | (.06) | .001 | .20 | (.06) | .001 | .18 | (.08) | .021 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.03 | (.05) | .471 | .03 | (.05) | .498 | -.02 | (.06) | .778 | .02 | (.05) | .719 | .01 | (.06) | .864 | .09 | (.06) | .152 |
| Cognitive ability | -.03 | (.02) | .135 | -.03 | (.03) | .251 | -.03 | (.03) | .178 | -.03 | (.03) | .334 | -.03 | (.03) | .322 | -.02 | (.03) | .512 |
| Conscientiousness | -.02 | (.03) | .406 | -.02 | (.03) | .405 | -.03 | (.03) | .353 | -.02 | (.03) | .434 | -.04 | (.03) | .205 | -.03 | (.03) | .421 |
| Test score | .06 | (.03) | .024 | .05 | (.03) | .144 | .06 | (.03) | .026 | .06 | (.03) | .053 | .07 | (.03) | .010 | .06 | (.03) | .064 |
| Self-concept | .00 | (.04) | .923 | .00 | (.05) | .991 | .00 | (.04) | .924 | -.02 | (.04) | .718 | -.01 | (.05) | .867 | -.03 | (.05) | .601 |
| HW self-efficacy | .01 | (.03) | .784 | .02 | (.03) | .447 | .01 | (.03) | .762 | .03 | (.03) | .311 | .02 | (.03) | .409 | .04 | (.04) | .232 |
| Intrinsic value | .03 | (.04) | .430 | .04 | (.03) | .227 | .02 | (.04) | .607 | .04 | (.03) | .226 | .03 | (.05) | .556 | .04 | (.04) | .250 |
| Utility value | .65 | (.03) | .000 | .59 | (.04) | .000 | .65 | (.03) | .000 | .60 | (.04) | .000 | .64 | (.03) | .000 | .57 | (.05) | .000 |
| Class' math valuing | .03 | (.03) | .295 | .08 | (.03) | .012 | .04 | (.04) | .211 | .08 | (.03) | .014 | .05 | (.04) | .201 | .06 | (.04) | .119 |
| Disruptions in class | -.04 | (.03) | .129 | -.04 | (.03) | .175 | -.05 | (.03) | .094 | -.04 | (.03) | .188 | -.06 | (.03) | .022 | -.04 | (.03) | .231 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | 8.67 | (1) | .003 | 4.75 | (1) | .029 | 10.20 | (1) | .001 | 10.80 | (1) | .001 | 4.75 | (1) | .029 | .82 | (1) | .366 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

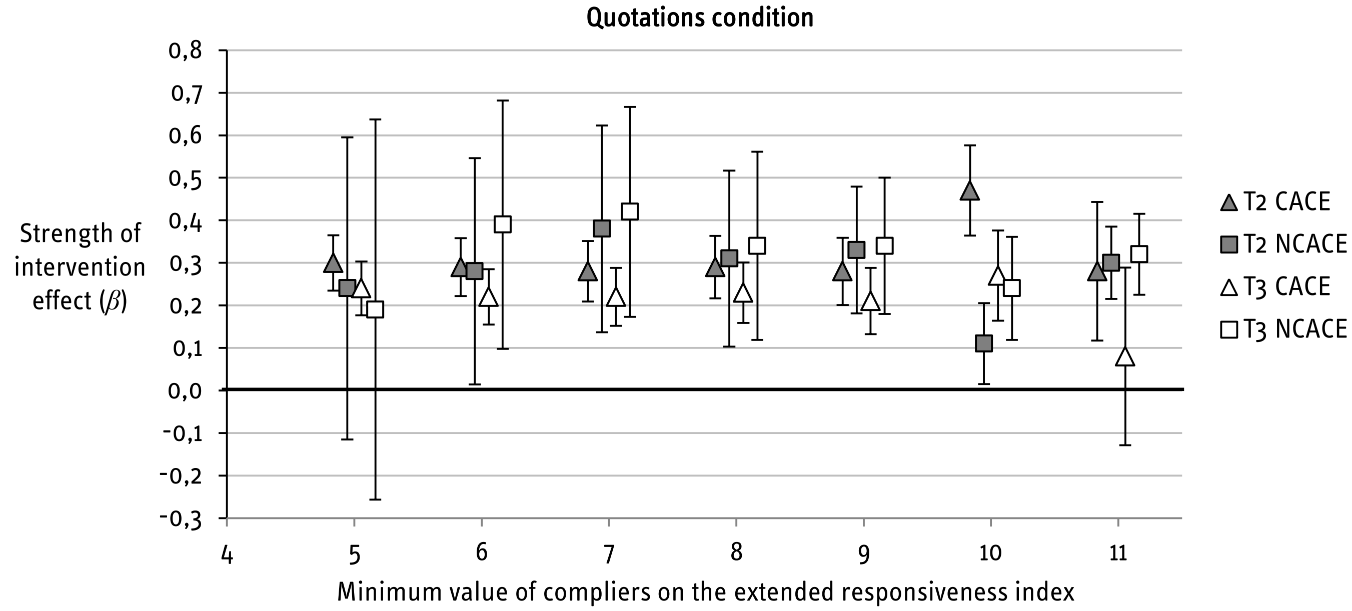
Table F8

*Intervention Effects on Students' Math Utility Value Beliefs in the Text Condition Depending on Students' Responsiveness: Cutoff on Extended Responsiveness Index above Values of 7 (NC Max. Value) and 8 (C Min. Value)*

| Responsiveness index | Cutoff | | | | | | Cutoff | | | | | | Cutoff | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC max. value | 8 | | | | | | 9 | | | | | | 10 | | | | | |
| C min. value | 9 | | | | | | 10 | | | | | | 11 | | | | | |
| Frequencies | *N* (%) | | | | | | *N* (%) | | | | | | *N* (%) | | | | | |
| Noncomplier | 276 (20%) | | | | | | 434 (32%) | | | | | | 784 (58%) | | | | | |
| Complier | 1079 (80%) | | | | | | 921 (68%) | | | | | | 571 (42%) | | | | | |
| Measurement point | T2 | | | T3 | | | T2 | | | T3 | | | T2 | | | T3 | | |
| Noncomplier | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* | $\beta_{NC}$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Text | -.13 | (.13) | .345 | .09 | (.20) | .648 | -.06 | (.10) | .588 | .14 | (.15) | .366 | .00 | (.07) | .971 | .17 | (.09) | .052 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | -.10 | (.10) | .306 | .08 | (.12) | .514 | -.08 | (.09) | .375 | .08 | (.10) | .451 | -.03 | (.07) | .703 | .04 | (.07) | .561 |
| Cognitive ability | -.09 | (.05) | .075 | -.12 | (.07) | .077 | -.05 | (.04) | .286 | -.09 | (.06) | .131 | -.02 | (.03) | .437 | -.05 | (.04) | .222 |
| Conscientiousness | -.06 | (.07) | .384 | -.10 | (.06) | .114 | -.04 | (.05) | .422 | -.11 | (.05) | .038 | -.04 | (.04) | .308 | -.07 | (.04) | .084 |
| Test score | .04 | (.05) | .464 | .04 | (.07) | .594 | .01 | (.05) | .750 | .00 | (.06) | .992 | -.02 | (.04) | .509 | -.03 | (.05) | .479 |
| Self-concept | -.03 | (.09) | .730 | .08 | (.11) | .464 | -.04 | (.08) | .618 | .06 | (.10) | .566 | -.04 | (.06) | .461 | -.02 | (.07) | .809 |
| HW self-efficacy | .10 | (.06) | .091 | .04 | (.09) | .614 | .08 | (.05) | .113 | .07 | (.07) | .365 | .06 | (.04) | .128 | .08 | (.05) | .116 |
| Intrinsic value | .04 | (.10) | .663 | -.06 | (.10) | .535 | .02 | (.08) | .772 | -.05 | (.08) | .542 | .06 | (.06) | .334 | .01 | (.07) | .944 |
| Utility value | .66 | (.09) | .000 | .49 | (.09) | .000 | .68 | (.07) | .000 | .52 | (.08) | .000 | .64 | (.05) | .000 | .48 | (.06) | .000 |
| Class' math valuing | -.01 | (.06) | .851 | .13 | (.08) | .113 | -.02 | (.05) | .741 | .12 | (.08) | .112 | .00 | (.04) | .940 | .10 | (.05) | .048 |
| Disruptions in class | .03 | (.06) | .637 | -.08 | (.07) | .227 | -.01 | (.06) | .821 | -.08 | (.06) | .197 | -.02 | (.04) | .611 | -.03 | (.04) | .429 |
| Complier | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* | $\beta_C$ | (SE) | *p* |
| Intervention | | | | | | | | | | | | | | | | | | |
| Quotations | .22 | (.07) | .002 | .15 | (.09) | .103 | .23 | (.07) | .002 | .13 | (.09) | .158 | .26 | (.09) | .002 | .09 | (.10) | .373 |
| Covariates | | | | | | | | | | | | | | | | | | |
| Gender (1 = male) | .04 | (.08) | .585 | .07 | (.06) | .255 | .06 | (.09) | .494 | .08 | (.07) | .268 | .02 | (.09) | .863 | .11 | (.07) | .133 |
| Cognitive ability | -.02 | (.03) | .629 | .00 | (.03) | .975 | -.04 | (.04) | .327 | -.01 | (.04) | .827 | -.10 | (.05) | .060 | -.07 | (.06) | .247 |
| Conscientiousness | -.03 | (.03) | .268 | -.03 | (.03) | .316 | -.05 | (.03) | .164 | -.03 | (.03) | .419 | -.05 | (.04) | .242 | -.06 | (.04) | .145 |
| Test score | .05 | (.03) | .117 | .05 | (.03) | .166 | .06 | (.03) | .059 | .06 | (.03) | .024 | .13 | (.04) | .002 | .11 | (.03) | .001 |
| Self-concept | .02 | (.05) | .722 | -.02 | (.06) | .741 | .03 | (.06) | .605 | -.01 | (.06) | .816 | .05 | (.08) | .505 | .06 | (.08) | .473 |
| HW self-efficacy | -.01 | (.03) | .839 | .01 | (.04) | .749 | -.01 | (.03) | .753 | -.01 | (.04) | .797 | -.02 | (.05) | .679 | -.05 | (.05) | .329 |
| Intrinsic value | .02 | (.05) | .760 | .04 | (.04) | .311 | .02 | (.05) | .684 | .05 | (.05) | .268 | .00 | (.07) | .950 | .02 | (.06) | .725 |
| Utility value | .63 | (.04) | .000 | .59 | (.04) | .000 | .62 | (.05) | .000 | .58 | (.05) | .000 | .66 | (.05) | .000 | .63 | (.05) | .000 |
| Class' math valuing | .05 | (.04) | .242 | .08 | (.04) | .028 | .06 | (.05) | .216 | .08 | (.04) | .032 | .08 | (.05) | .143 | .11 | (.05) | .028 |
| Disruptions in class | -.05 | (.03) | .106 | -.01 | (.03) | .788 | -.04 | (.03) | .285 | .00 | (.04) | .967 | -.05 | (.04) | .227 | -.03 | (.03) | .319 |
| Wald-$X^2$-test | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* | $X^2$ | *(df)* | *p* |
| $\beta_{NC} = \beta_C$ | 4.36 | (1) | .037 | .04 | (1) | .833 | 4.31 | (1) | .038 | .00 | (1) | .967 | 5.20 | (1) | .023 | .34 | (1) | .561 |

*Notes.* T = time; NC = noncomplier; C = complier; max. = maximum; min. = minimum; *N* = number; HW = homework; $\beta$ = standardized regression coefficient; *SE* = standard error; *p* = *p*-value; *df* = degrees of freedom.

**Graphic chart of intervention effects depending on students' degree of intervention responsiveness (extended responsiveness index including the intervention boosters)**



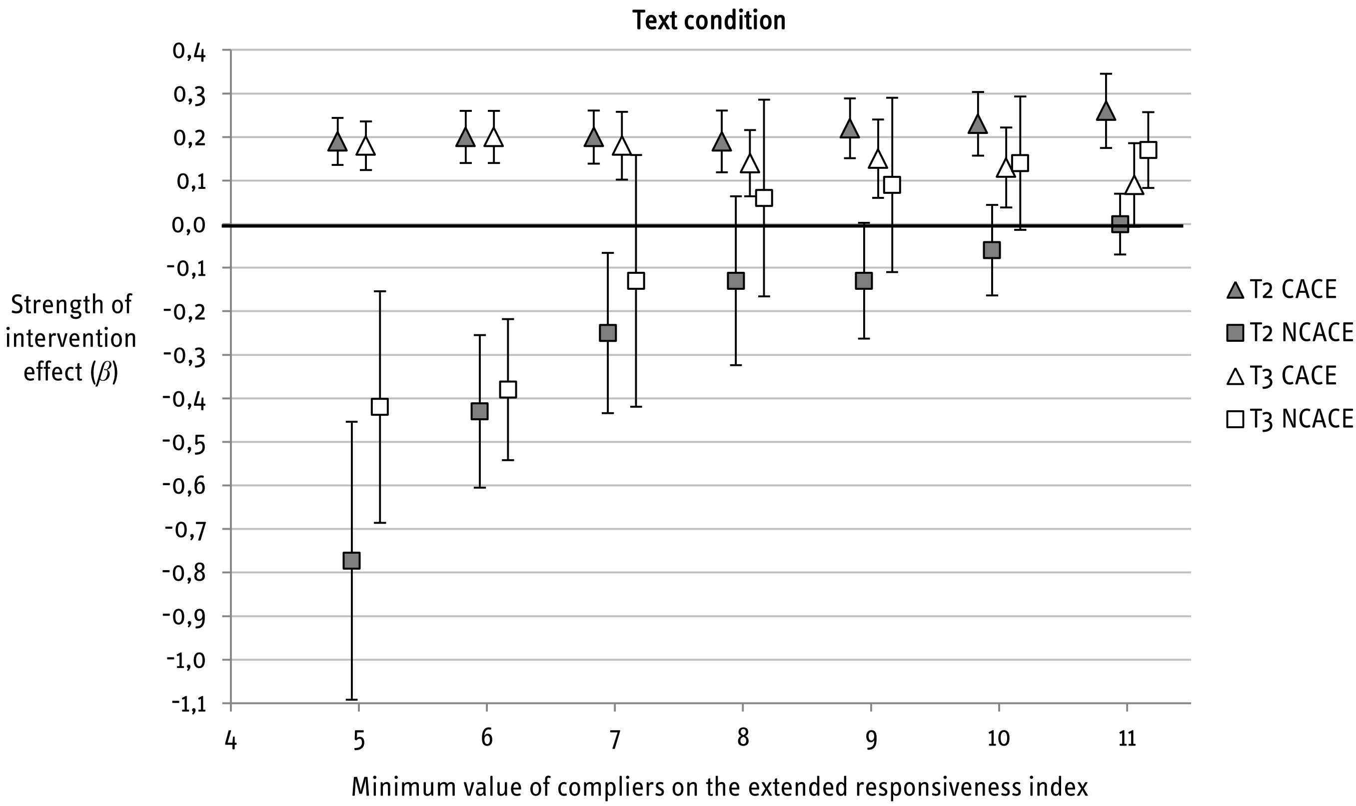




**Figure F2:** Effects of the quotations condition (above) and of the text condition (below) on students' utility value beliefs for different cutoff values depending on students' intervention responsiveness (responsiveness index + number of completed boosters), with standard error bars.

**Part G) Utility value beliefs as a mediator of intervention effects on competence beliefs**

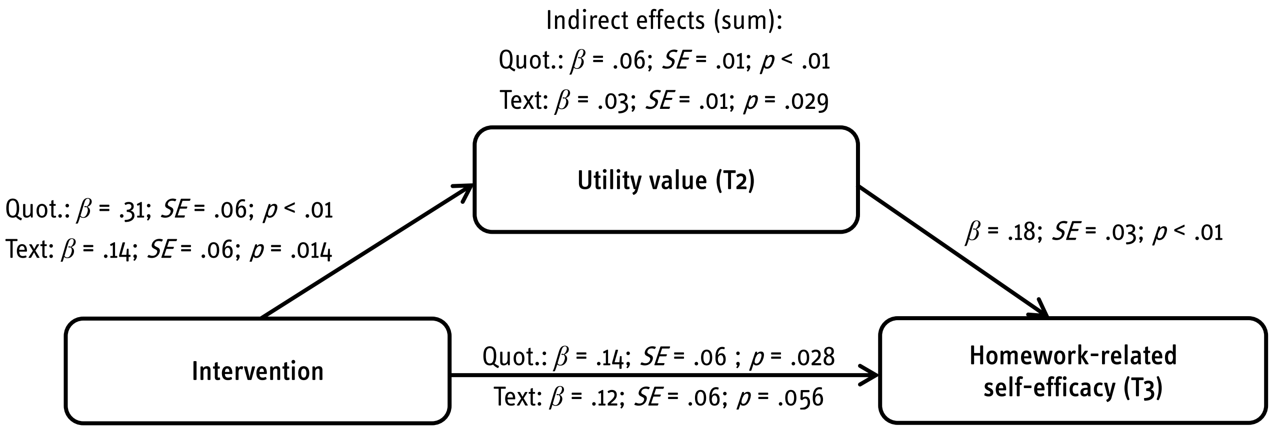



**Figure G1:** Mediation of intervention effects on students' self-efficacy through students' utility value beliefs.

*Notes*. Quot. = quotations condition; Text = text condition; $\beta$ = standardized regression coefficient; $p$ = two-tailed $p$-value; T2 = posttest (six weeks after the intervention); T3 = follow-up test (five months after the intervention). Solid paths represent effects significant at $p < .06$.

**Information on the statistical analysis**

The mediation model was calculated in Mplus (Version 8; Muthén & Muthén, 1998–2012) following the simple mediation model described in Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209-233. doi: 10.1037/a0020141. The mediator was regressed on two dummy variables indicating the intervention conditions (quotations, text), controlling for the intial value of the mediator. The outcome was regressed on the two dummies representing the intervention conditions (quotations, text) and on the mediator, controlling for the intial value of the outcome.

Standard errors were corrected to account for the nesting of students within classes by using the TYPE IS COMPLEX function integrated in Mplus (i.e., design-based correction of standard errors and test statistics; cf., McNeish et al., 2017).

The full information maximum likelihood method integrated in Mplus was used to deal with missing data (e.g., Graham, 2009). To make the assumption of missing at random more plausible, correlations of three auxiliary variables (students' gender; cognitive ability score; end-of-year math grade in Grade 8) with the predictors and the outcome were included in the model (e.g., Enders, 2010). The auxiliaries' and predictors' residuals were also included in the model.

All continuous variables were standardized prior to the analyses. Consequently, the regression coefficients can be directly interpreted as measures of effect sizes, for example of the interventions on the mediator as compared to the control condition (e.g., Tymms, 2004).