AERA Open
April-June 2016, Vol. 2, No. 2, pp. 1–14
DOI: 10.1177/2332858416649168
© The Author(s) 2016. http://ero.sagepub.com

Side Effects of Motivational Interventions? Effects of an Intervention in Math Classrooms on Motivation in Verbal Domains

Hanna Gaspard
Hector Research Institute of Education Sciences and Psychology, University of Tübingen
Anna-Lena Dicke
University of California, Irvine
Barbara Flunger
Isabelle Häfner
Brigitte M. Brisson
Ulrich Trautwein
Benjamin Nagengast
Hector Research Institute of Education Sciences and Psychology, University of Tübingen

One way to address the leaking pipeline toward STEM-related careers (i.e., science, technology, engineering, and mathematics) is to intervene on students' STEM motivation in school. However, a neglected question in intervention research is how such interventions affect motivation in subjects not targeted by the intervention. This question was addressed through data from a cluster-randomized study in which a value intervention was successfully implemented in 82 ninth-grade math classrooms. Side effects on value, self-concept, and effort in German as students' native language and English as a foreign language were assessed 6 weeks and 5 months after the intervention. Negative effects on value for German, but not for English, were found 5 months after the intervention. The theoretical and educational implications of such effects are discussed.

Keywords: *expectancy-value theory, dimensional comparisons, motivational intervention, self-concept, value*

In many Western countries, concerns have been raised about a lack of young people choosing careers in science, technology, engineering, and mathematics (STEM; e.g., National Science Board, 2007). Important precursors of career choices are high school students' motivational beliefs about their expectancies and values for different subjects (Eccles et al., 1983; for a review, see Wang & Degol, 2013). One possible way to address the leaking pipeline toward STEM careers at an early stage is thus to foster motivation for related subjects, such as math, in high school. Researchers have recently developed a number of successful motivational interventions in STEM (for an overview, see Karabenick & Urdan, 2014). Some of these draw on expectancy-value theory (Eccles et al., 1983) and aim at helping students understand the value of STEM courses. Previous studies have shown that value interventions can be effective in promoting motivation and performance in STEM courses as well as STEM course choices (Harackiewicz, Rozek, Hulleman, & Hyde, 2012; Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009).

However, intervening on student motivation in the classroom means intervening in a complex system, which might bring side effects with it. While implementation science calls for regularly assessing side effects of interventions (see Moher, Schulz, & Altman, 2001), intervention research in the area of motivation has so far neglected potential side effects. Particularly, previous intervention studies in STEM did not examine potential effects on motivation in non-STEM areas. Students' expectancies and values are highly domain specific (Bong, 2001). Students tend to see themselves as either mathematically or verbally oriented, irrespective of whether their achievement in these domains differs substantially (Marsh & Hau, 2004). Academic choices, in turn, are influenced by intraindividual hierarchies in motivational beliefs: The probability that a student intends to pursue a career in STEM increases not only with his or her motivation in STEM becoming higher but also with his or her motivation in other domains becoming lower (Chow, Eccles, & Salmela-Aro, 2012; Eccles, 2009; Parker et al., 2012). What happens to these motivational patterns

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License (http://www.creativecommons.org/licenses/by/3.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

when motivation in one domain is fostered through interventions? Expectancy-value theory (Eccles et al., 1983) and dimensional comparison theory (Möller & Marsh, 2013) implicitly assume that increased motivation in one domain can have negative effects on dissimilar domains. Interventions targeting STEM motivation are a possibility to investigate this assumption experimentally. In this study, we examine possible negative side effects of a value intervention in mathematics classrooms on students' motivation for their native language and a foreign language.

Student Motivation and Dimensional Comparisons

According to expectancy-value theory (Eccles et al., 1983), academic choices, such as choosing a university major, are made on the basis of two beliefs: (a) the expectancy that one can succeed in a task and (b) the value that one attaches to a task. Expectancies are closely related to academic self-concepts, referring to students' evaluation of their abilities in a given domain (Bong & Skaalvik, 2003; Eccles & Wigfield, 2002). Task value comprises several components: attainment value or the personal importance to do well, intrinsic value or enjoyment, utility value or the usefulness for personal goals, and cost or the perceived negative aspects of engaging in a task. Previous research has found high correlations among these components, and many studies collapsed the positive value aspects into a single scale (Trautwein et al., 2013). Furthermore, expectancies and values have been found to be closely related, and this association increases with students' age (Wigfield, Tonks, & Klauda, 2009). An extensive line of research demonstrates that expectancies and values are indeed important predictors for achievement-related behaviors, such as effort, as well as for academic choices (for reviews, see Wang & Degol, 2013; Wigfield et al., 2009).

Expectancies and values are developed through experiences with different domains in the school context. These experiences provide students with a set of possible comparisons, including other students' achievement, but also comparisons among domains. Dimensional comparison theory (Möller & Marsh, 2013) focuses on comparisons among domains. It assumes that individuals compare their ability in one domain with their ability in another domain (e.g., "How good am I in English compared with math?"). Reasons why students engage in such dimensional comparisons may include self-evaluation in terms of knowing one's strengths and weaknesses, as well as self-enhancement, which can be achieved through upward comparisons with a more favorable domain (Möller & Husemann, 2006; Möller & Marsh, 2013). In educational psychology, research has mainly investigated dimensional comparisons in the context of self-concept. Much of this research was based on the internal/external frame-of-reference model (Marsh, 1986), which describes the association between self-concept and achievement in different domains, particularly math and verbal domains. Supporting this model, path-analytic studies have found that achievement in one domain (e.g., math) can have negative effects on self-concept in another domain (e.g., English; Marsh & Hau, 2004; Möller, Pohlmann, Köller, & Marsh, 2009). Such contrast effects have mainly been supported for comparisons between math and the native language but also for other comparisons between numerical (e.g., physics) and verbal domains (e.g., foreign language; Marsh et al., 2015). These contrast effects seem to depend on students' beliefs about the dissimilarity of these subjects, with stronger contrast effects being found for students who believe math and verbal abilities to be inversely related (Möller, Streblow, & Pohlmann, 2006). Beyond self-concept, effects of dimensional comparisons were found on interest (Schurtz, Pfost, Nagengast, & Artelt, 2014) and enjoyment (Goetz, Frenzel, Hall, & Pekrun, 2008). Whereas there is extensive evidence on the role of dimensional comparisons stemming from correlational research, there are fewer experimental studies, which do, however, support the assumptions of dimensional comparison theory (see Meyer, 1982; Möller & Köller, 2001; Pohlmann & Möller, 2009, Study 3).

Dimensional comparisons also play a crucial role in expectancy-value theory. Academic choices are supposed to be informed by intraindividual hierarchies of expectancies and values (Eccles, 2009). Recent research has addressed this assumption showing that choices (e.g., beginning a math- vs. verbal-intensive major) are affected not only by expectancies and values in the target domain but also by expectancies and values in other domains (Chow et al., 2012; Nagy et al., 2008; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006; Parker et al., 2012). Eccles (2009) proposed that such intraindividual hierarchies result from individuals comparing their performance and the effort they need to succeed across domains. Put simply, students have limited time and energy and cannot engage in all subjects to the same extent. The lower their expectancies and values in one subject compared with other subjects, the lower the probability that they will put much effort into engaging in this particular subject (Trautwein & Lüdtke, 2007). Expectancy-value theory thus assumes not only expectancies but also values to be affected by dimensional comparison processes. When these results are transferred to intervention research, it could be that highlighting the value of one subject can lead to changes in students' hierarchies of importance and increase the subjective costs of engaging in another subject. Given the findings on dimensional comparisons, we propose that motivational interventions in STEM can have adverse effects on motivation in verbal domains.

Value Interventions

One intervention approach that has been applied to foster STEM motivation is value interventions (Harackiewicz,

Tibbetts, Canning, & Hyde, 2014). Drawing on expectancy-value theory (Eccles et al., 1983), these interventions try to foster domain-specific motivation by targeting one component of task value: utility value. Through stimulating the perceived relevance of the learning material for students' lives, these interventions ultimately aim at reinforcing their interest and engagement in the STEM domain (Harackiewicz et al., 2014). In line with this assumption, value interventions have been found to foster not only utility value but also interest and course choices (Harackiewicz et al., 2012; Hulleman et al., 2010; Hulleman & Harackiewicz, 2009). Other motivational outcomes, such as expectancies and effort, have not been regularly considered. However, when learning a math technique in the laboratory, Western students who received information on its short-term usefulness benefited in terms of not only their interest for this technique but also their effort and perceived competence (Shechter, Durik, Miyamoto, & Harackiewicz, 2011). Different strategies have been used to foster perceived relevance to date, which—among others—differ in that students were either presented with arguments for the relevance of the learning material or encouraged to generate arguments for its relevance themselves. It is still unclear which intervention strategies are most effective in fostering student motivation. This also seems to depend on students' characteristics: Whereas students with high initial motivation tended to benefit most from information on the relevance of the learning material, students with lower initial motivation tended to benefit more from being encouraged to self-generate utility arguments (Durik, Hulleman, & Harackiewicz, 2015).

Present Investigation

In the present investigation, we test whether a motivational intervention in math had negative side effects on motivation for verbal subjects using data from a large cluster-randomized trial conducted in ninth-grade classrooms in Germany (see Brisson et al., 2016; Gaspard, Dicke, Flunger, Brisson, et al., 2015). Motivation in math was chosen as the target of the intervention, as high school math courses are one important prerequisite for future careers in STEM fields. In this intervention study, 82 classrooms were randomly assigned to one of two intervention conditions or a waiting control group. Drawing on expectancy-value theory, the intervention consisted of a 90-minute session in a math classroom focusing on the value of math for students' lives. Students in the two intervention conditions were first presented with information on the relevance of mathematics and afterward worked on different kinds of relevance-inducing tasks, which differed in the extent to which arguments had to be self-generated. Students in the quotations condition evaluated interview quotations describing the usefulness of math, whereas students in the text condition wrote an essay on the relevance of math. Whereas both intervention conditions were shown to positively affect utility value, the quotations condition additionally showed positive effects on attainment and intrinsic value, as well as on self-concept and achievement (Brisson et al., 2016; Gaspard, Dicke, Flunger, Brisson, et al., 2015).

Here, we test effects of this intervention on the patterns of motivation across several domains. Our major aim was to explore intervention effects on motivation in German as students' native language and English as students' first foreign language. As a means of comparison, we also report the respective effects on motivation in math as the target domain. To examine the breadth of intervention effects, we consider effects the most proximal outcome, value, and two more distal outcomes, self-concept and effort. Given the literature on dimensional comparisons, we expected any side effects on motivation in verbal subjects to be negative. Because side effects rely on dimensional comparison processes and because effects across domains are typically smaller than effects within domains (Möller et al., 2009), we expected to find the following pattern of results. First, in comparing interventions, stronger side effects were expected for interventions that show stronger effects on motivation in the target subject. With respect to our study, the quotations condition was shown to be more successful in promoting motivation in math (Brisson et al., 2016; Gaspard, Dicke, Flunger, Brisson, et al., 2015) and therefore was expected to also show stronger side effects. Second, regarding different outcomes, stronger side effects should be found for those motivational variables that are more effectively promoted in the target subject. Whereas the intervention directly targeted students' value beliefs, there is only some first evidence that value interventions can also be effective in fostering students' expectancy-related beliefs and effort (Brisson et al., 2016; Shechter et al., 2011). Therefore, in line with the focus of the intervention, stronger side effects were expected for value beliefs as compared with self-concept and effort. Third, regarding different subjects, contrast effects are typically found for students' native language, whereas there is less evidence for foreign languages (Möller & Marsh, 2013). We therefore expected German to be the main target of dimensional comparison processes. Also, given the nature of the intervention, the perceived relevance of these subjects might be important. Students in Germany tend to perceive German as less relevant than English for their current and future lives (see Goetz et al., 2014), and this might lead to more negative side effects for German as compared with English.

Method

Sample

Data were collected in 82 ninth-grade classes in 25 academic track schools in the German state of Baden-Württemberg. The sample size was based on a power

analysis for a multisite cluster-randomized trial aiming at an acceptable power ($\beta > .70$) to detect intervention effects ($\delta = 0.20$) when comparing a single intervention condition with the control condition (for more details, see Gaspard, Dicke, Flunger, Brisson, et al., 2015). A total of 1,978 students with active parental consent participated in the study, corresponding to a 96% participation rate. For the current study, 62 students in the two intervention conditions were excluded, as they were absent during the intervention. Data analyses were thus based on a sample of 1,916 students (age at the beginning of the study: M = 14.62, SD = 0.47; 53.5% female). The study consisted of three waves of data collection from September 2012 to March 2013. Students were administered questionnaires by trained research assistants before the intervention (pretest = T1), 6 weeks after the intervention (posttest = T2), and 5 months after the intervention (follow-up = T3).

Value Intervention in Math

Before the first data collection, within each school, the participating teachers (n = 73) and their classes (n = 82) were randomly assigned to one of two intervention conditions or a waiting control condition. For teachers participating with two classes (n = 9), both classes were assigned to the same experimental condition to avoid diffusion effects. This assignment resulted in unequal class sample sizes for different conditions (quotations condition: 25 classes; text condition: 30 classes; waiting control condition: 27 classes).

Students in the intervention conditions received a 90-minute standardized value intervention led by five trained researchers. The intervention consisted of a psychoeducational presentation on the relevance of math for the whole class and tasks for individual students. The psychoeducational presentation had two main components. First, research results on the importance of effort and self-concept for math achievement were presented. Students were also told about frame-of-reference effects (i.e., effects of social comparisons in the classroom) and the benefits of relying on temporal instead of social comparisons. This first part aimed at inoculating students against potential negative effects of highlighting the importance of a subject. These might occur if students judge their own achievement in this subject as low and are therefore threatened by information on its importance (cf., Durik, Shechter, Noh, Rozek, & Harackiewicz, 2015). Second, students were provided with various examples of the relevance of math for future education, career opportunities, and leisure time activities. This presentation was identical for both intervention conditions. After this presentation, students worked on individual tasks, which differed between the two conditions. In the quotations condition, students were asked to read quotations of young adults describing situations in which math was useful to them and to evaluate these quotations based on their personal relevance. In the text condition, students were asked to make a list of arguments for the personal relevance of math to their current and future lives and to write an essay explaining these arguments. Thus, in both conditions, the students had to apply the relevance of mathematics to their lives, whereas the two conditions differed in the specific structure of the task and the extent to which arguments had to be self-generated.

Additionally, each intervention group received two booster tasks that were embedded into a homework diary, which was filled out by all classes for 4 weeks after the intervention. The first booster task, in which students were asked to reproduce what they remembered from their individual tasks, was filled out 1 week after the intervention. The second booster task was filled out 2 weeks after the intervention and resembled the individual tasks assigned to the students during the intervention lesson (for more details on the intervention, see Gaspard, Dicke, Flunger, Brisson, et al., 2015). Classes in the waiting control condition also filled out homework diaries, but these did not include any booster tasks. Students in the waiting control condition received the intervention that was shown to be more successful after the last wave of data collection.

The intervention focused only on the subject of math. No explicit comparisons between the importance of math and other subjects were made during the intervention. All the intervention materials referred to the relevance of math. Students were told that this session was delivered in math classrooms because math is a subject that is often experienced as difficult and useless. If students asked questions referring to other subjects (e.g., "What about the relevance of other subjects?"), the researcher conducting the intervention said that other subjects were relevant as well but that the focus of this lesson was on the relevance of math.

Measures

We assessed value beliefs, self-concept, and effort for math, German, and English with parallel scales (i.e., the wording was identical except for the subject name). All items are reported in the online supplement. As a response format, a 4-point Likert scale ranging from *completely disagree* to *completely agree* was used for all items.

Value beliefs. Value beliefs were assessed with four items for each subject. The items tapped different value aspects: attainment value, intrinsic value, and utility value. The scales for German and English were constructed with a subset of items out of a larger questionnaire assessing value beliefs in math (Gaspard, Dicke, Flunger, Schreier, et al., 2015). According to preliminary factor analyses, we excluded one item assessing cost ("[Subject] is a real burden to me."). All resulting scales exhibited good internal consistency (α at T1, T2, T3, respectively: math = .77, .78, .77; German = .85, .85, .86; English = .83, .84, .84).

Self-concept. Self-concept was assessed with four items. All items were well validated stemming from previous German large-scale studies (e.g., Trautwein, Lüdtke, Köller, & Baumert, 2006). The internal consistency of this scale was good for all subjects at all measurement waves (α at T1, T2, T3, respectively: math = .93, .92, .92; German = .89, .89, .90; English = .90, .90, .91).

Effort. Effort in the subjects math, German, and English was assessed with four items for each subject (adapted from Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009). The scale showed good internal consistency for all subjects (α at T1, T2, T3, respectively: math = .80, .84, .86; German = .88, .89, .89; English = .85, .87, .88).

As a prerequisite for our analyses, we conducted tests of measurement invariance across time, subjects, and intervention conditions in several steps (see Tables S1–S3 in the supplemental material for more details). Specifically, we tested invariance of factor loadings (strict measurement invariance) and invariance of item intercepts (strong measurement invariance) to compare differences in latent means (Widaman & Reise, 1997). In the first step, we conducted tests of measurement invariance for value, self-concept, and effort across the three time points. As the analyses suggested that strong measurement invariance across time was acceptable for all three constructs, we used these models constraining factor loadings and intercepts to be equal across time for further tests of measurement invariance across subjects and intervention conditions. For the tests across subjects, a model with equal intercepts (in addition to factor loadings) was not tenable for value. We therefore assessed partial strong measurement invariance (Steenkamp & Baumgartner, 2009) by freely estimating the intercept for one item (assessing utility value). As this model yielded an acceptable fit, partial strong measurement invariance was defensible for value. For effort and self-concept, the tests of measurement invariance across subjects did not suggest any problem. The tests across intervention conditions indicated that strong measurement invariance could be accepted for all three constructs. Comparability of latent means across time, subjects, and intervention conditions was therefore established.

Statistical Analyses

Multilevel structural equation modeling. Given the multilevel structure of the data, we used multilevel structural equation modeling with Mplus 7 (Muthén & Muthén, 1998–2012) to examine the effects of the intervention on students' value beliefs, self-concept, and effort. Multilevel structural equation modeling (Mehta & Neale, 2005) combines the advantages of multilevel analyses, which take the nesting of students in classrooms into account (Raudenbush & Bryk, 2002), and latent variable modeling, which controls for measurement error (Bollen, 1989). An additional advantage of structural equation modeling is its flexibility; for instance, it allows explicit modeling of the measurement properties that were established according to prior confirmatory factor analyses.

Multilevel structural equation analyses were carried out separately for the latent constructs value, self-concept, and effort at the posttest and follow-up (for the estimated model, see Figure 1). We combined all three subjects into one model for each construct and time point.[1] In line with the recommendations for the evaluation of cluster randomized trials (Raudenbush, 1997), the respective pretest constructs in all three subjects were included as control variables at the student level as well as the class level. The effects at both levels were freely estimated to account for contextual effects (Marsh et al., 2009). The indicators of the latent constructs at the student level were group-mean centered, and manifest aggregation was used for the class-level indicators (Marsh et al., 2009). Factor loadings were set to be equal across levels to ensure a common metric at student and class levels (Marsh et al., 2009). Additionally, the factor loadings and item intercepts were constrained to be the same across time and subjects (with the exception of one value item, for which the intercept was freely estimated across subjects; see above). To assess effects of the intervention, we regressed the latent constructs at the posttest/follow-up on two class-level dummy variables that indicated the intervention conditions (quotations, text) with the control condition as a reference group.

To obtain standardized effect sizes (for effect sizes in multilevel models, see Marsh et al., 2009; Tymms, 2004), the raw coefficients of intervention effects were divided by the total variance of the outcome variables out of null models (i.e., allowing all variables to correlate instead of estimating path coefficients). These effect sizes thus represent the adjusted difference between the intervention conditions and the control condition in the outcome variable in total standard deviations.

Missing data. Due to the absence of students at single-measurement waves and nonresponse to single items, missing data ranged from 5.4% to 12.6% for the indicators of the focal motivational constructs. All analyses were conducted with full information maximum likelihood estimation (Graham, 2009) implemented in Mplus. To make the assumption of missing at random more plausible (see Enders, 2010), a nonverbal cognitive ability score, gender, previous math grades, and achievement data for math at T1 were used as auxiliary variables. To incorporate these auxiliary variables in our model, we used a saturated correlates model (Graham, 2003) correlating these variables with one another as well as with the manifest predictor variables (i.e., the dummy variables for the intervention condition) and the residuals of the indicator variables for all latent constructs at both levels.

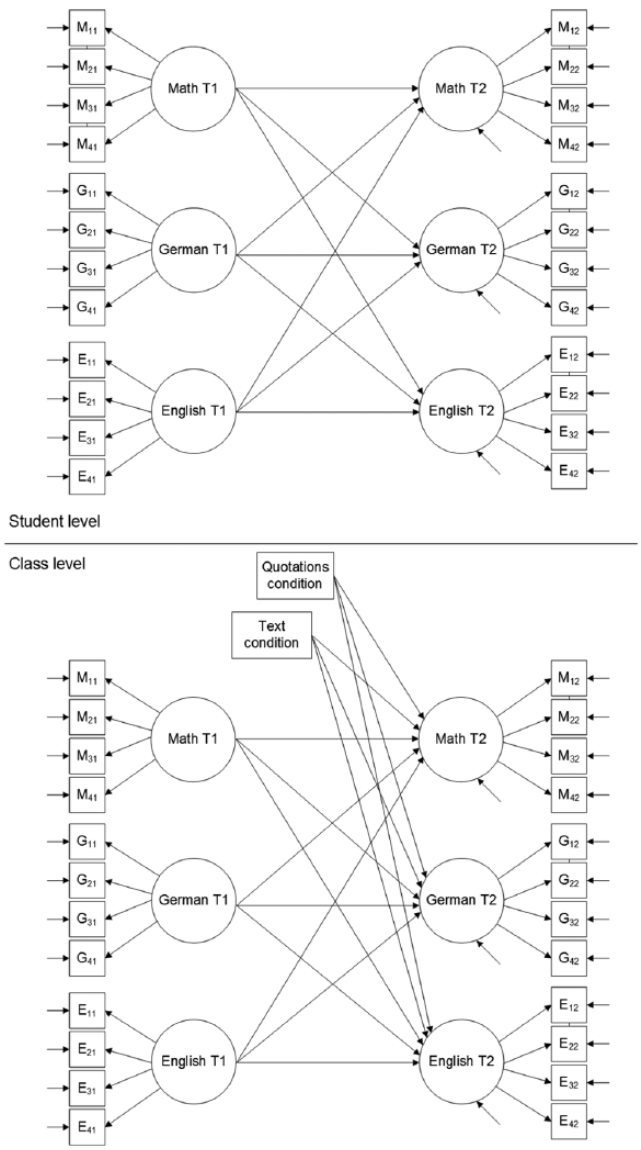

FIGURE 1. *Multilevel structural equation modeling to estimate intervention effects on value, self-concept, and effort in math, German, and English. Covariances between predictor variables at both levels as well as between residual variances of identical item stems at the within level are not depicted. The model additionally included auxiliary variables, which were correlated with the dummy variables indicating the two intervention conditions and the residuals of the indicator variables for all latent constructs at both levels.*

Results

Descriptive Statistics and Randomization Check

Descriptive statistics for all scales are displayed in Table 1. Several aspects can be observed by inspecting the means. First, students in all conditions reported relatively high levels of value, self-concept, and effort in English as compared with math and German. Second, mean levels were relatively stable across the three measurements. In the control condition, which can be seen as a comparison standard, the motivation for math tended to decrease; motivation for German tended to increase; and motivation for English remained relatively high.

Correlations among value, self-concept, and effort in all subjects from a confirmatory factor analysis with the pretest data are presented in Table 2. The confirmatory factor analysis supported the separability of the three constructs across the three subjects (χ^2 = 2663.12, *df* = 513, comparative fit index = .940, Tucker-Lewis index = .926, root mean square error of approximation = .048, standardized root mean square residual = .051). Several aspects can be noted with regard to the correlation pattern. First, value, self-concept, and effort within one subject showed moderate to high intercorrelations. Second, value, self-concept, and effort between German and English showed low to moderate intercorrelations. Third, value and self-concept between math and the verbal domains tended to be negatively correlated. Fourth, the correlation pattern for effort indicated a lower degree of domain specificity with moderate positive intercorrelations among all three subjects.

To test if there were any differences among the three experimental conditions before the intervention, we specified multilevel, multigroup models (with each experimental condition as a group) for initial value, self-concept, and effort in math, German, and English as well as for the auxiliary variables (i.e., cognitive abilities, gender composition, math grades, math achievement test; see Table S4 in the Supplemental Material for details). We conducted omnibus tests comparing the means of the three groups by Wald χ^2 tests (using the "model test" command in Mplus), which are asymptotically equivalent to likelihood ratio tests (cf. Bollen, 1989). We found no significant differences among the groups prior to the intervention, neither for the focal constructs (all *p*'s ≥ .620) nor for the auxiliary variables (all *p*'s ≥ .125).

Intervention Effects on Value, Self-Concept, and Effort in Math, German, and English

Effects of the two intervention conditions as compared with the control condition were assessed on value, self-concept, and effort in math, German, and English at posttest and follow-up (see Table 3 for effect sizes and Table S5 in the Supplemental Material for the complete models, including the effects of pretest variables). In these analyses, we controlled for the respective construct at the pretest in all three subjects to achieve more precise estimates of the intervention effects (Raudenbush, 1997). Effects on value and self-concept in math as the target subject of the intervention have already been reported by Gaspard, Dicke, Flunger, Brisson, et al. (2015) and Brisson et al. (2016). For value, previous analyses used a more differentiated measure with subscales

TABLE 1
Descriptive Statistics for All Study Variables by Intervention Condition

	Quotations Condition			Text Condition			Control Condition		
Variable	*M*	*SD*	ICC	*M*	*SD*	ICC	*M*	*SD*	ICC
Math									
Value T1	2.68	0.64	.02	2.65	0.62	.07	2.61	0.60	.07
Value T2	2.74	0.63	.04	2.61	0.64	.07	2.55	0.64	.05
Value T3	2.75	0.61	.03	2.67	0.61	.08	2.58	0.61	.07
Self-concept T1	2.71	0.84	.02	2.68	0.86	.04	2.61	0.86	.04
Self-concept T2	2.75	0.84	.02	2.68	0.84	.05	2.59	0.84	.06
Self-concept T3	2.80	0.79	.01	2.75	0.79	.05	2.67	0.79	.05
Effort T1	2.78	0.59	.05	2.77	0.62	.01	2.78	0.61	.03
Effort T2	2.79	0.65	.07	2.73	0.67	.03	2.75	0.69	.02
Effort T3	2.71	0.68	.05	2.70	0.71	.04	2.68	0.69	.04
German									
Value T1	2.76	0.73	.08	2.73	0.71	.06	2.78	0.69	.04
Value T2	2.81	0.75	.10	2.81	0.75	.06	2.87	0.72	.05
Value T3	2.79	0.73	.07	2.87	0.74	.07	2.91	0.72	.07
Self-concept T1	2.96	0.75	.08	3.01	0.72	.07	2.99	0.73	.03
Self-concept T2	2.95	0.77	.09	3.02	0.73	.06	3.02	0.73	.02
Self-concept T3	2.91	0.78	.07	3.03	0.73	.08	3.00	0.73	.02
Effort T1	2.78	0.67	.06	2.81	0.64	.04	2.82	0.65	.06
Effort T2	2.83	0.71	.06	2.83	0.72	.04	2.85	0.70	.03
Effort T3	2.80	0.69	.06	2.84	0.73	.07	2.90	0.70	.04
English									
Value T1	3.35	0.57	.08	3.34	0.61	.06	3.33	0.58	.05
Value T2	3.37	0.58	.06	3.30	0.66	.04	3.33	0.61	.04
Value T3	3.38	0.60	.07	3.35	0.61	.06	3.31	0.62	.05
Self-concept T1	3.16	0.73	.04	3.14	0.70	.03	3.13	0.74	.07
Self-concept T2	3.17	0.70	.03	3.10	0.76	.02	3.12	0.72	.08
Self-concept T3	3.20	0.72	.04	3.13	0.73	.04	3.12	0.73	.06
Effort T1	3.06	0.61	.03	3.09	0.63	.06	3.13	0.60	.06
Effort T2	3.13	0.63	.03	3.09	0.69	.04	3.09	0.66	.06
Effort T3	3.06	0.65	.01	3.09	0.69	.06	3.08	0.69	.10

Note. Due to the absence of students at single-measurement waves and nonresponse to single items, the sample sizes for the scales range from 509 to 530 in the quotations condition, 619 to 680 in the text condition, and 550 to 609 in the control condition. ICC = intraclass correlation coefficient; T1 = pretest; T2 = posttest; T3 = follow-up.

for different value components. Here, the results on math as the target subject of the intervention are reported as a comparison to effects on German and English using parallel scales. All effects of the two intervention conditions reported in the text are standardized effect sizes with respect to the total variance of the outcome variable.

For value at the posttest, students in classes in the quotations condition reported higher math value when compared with students in classes in the control condition, controlling for their initial value in math, German, and English, $\beta = .28$, $p < .001$, 95% confidence interval (95% CI) [0.15, 0.41]. The quotations condition did not show a significant effect on German ($p = .104$) or English ($p = .940$) value at the posttest. At the follow-up, we still observed a positive effect of the quotations condition on math value, $\beta = .26$, $p < .001$, 95% CI [0.12, 0.39]. In addition, we found that students in classes in the quotations condition reported lower German value than students in classes in the control condition while controlling for students' initial values in all three subjects, $\beta = -.18$, $p = .004$, 95% CI [−0.30, −0.06]. No significant effect of the quotations condition was found on English value ($p = .090$) at the follow-up. The text condition did not have any significant effect on students' value beliefs at the posttest (all p's $\geq .077$) and the follow-up (all p's $\geq .117$). The effects of the two intervention conditions on students' value beliefs are further displayed in Figure 2.

For self-concept, students in classes in the quotations condition reported higher math self-concept when compared

TABLE 2
Correlations Between Study Variables at T1: Corrected for Measurement Error

Variable	1	2	3	4	5	6	7	8	9
1. Math value	—								
2. Math self-concept	.79***	—							
3. Math effort	.54***	.28***	—						
4. German value	−.06	−.27***	.17***	—					
5. German self-concept	−.25***	−.27***	.02	.73***	—				
6. German effort	.08*	−.13***	.46***	.77***	.49***	—			
7. English value	−.10**	−.22***	.09***	.31***	.30***	.23***	—		
8. English self-concept	−.13***	−.18***	.07*	.21***	.31***	.17***	.86***	—	
9. English effort	.05	−.13***	.44***	.30***	.19***	.56***	.66***	.53***	—

Note. Bivariate correlations from confirmatory factor analyses using pretest data are presented. Correlation pattern at T2 and T3 are comparable. T1 = pretest; T2 = posttest; T3 = follow-up.
*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 3
Effects of the Intervention Conditions as Compared With the Control Condition

	Quotations Condition						Text Condition					
	Posttest			Follow-Up			Posttest			Follow-Up		
Outcome	β	*p*	95% CI	β	*p*	95% CI	β	*p*	95% CI	β	*p*	95% CI
Value												
Math	0.28	<.001	[0.15, 0.41]	0.26	<.001	[0.12, 0.39]	0.05	.513	[−0.09, 0.18]	0.14	.062	[−0.01, 0.28]
German	−0.08	.104	[−0.19, 0.02]	−0.18	.004	[−0.30, −0.06]	−0.07	.187	[−0.17, 0.03]	−0.02	.782	[−0.15, 0.11]
English	0.00	.940	[−0.10, 0.11]	0.12	.090	[−0.02, 0.26]	−0.08	.077	[−0.18, 0.01]	0.10	.117	[−0.02, 0.22]
Self-concept												
Math	0.10	.049	[0.00, 0.20]	0.07	.180	[−0.04, 0.17]	0.04	.446	[−0.06, 0.13]	0.03	.510	[−0.06, 0.12]
German	−0.06	.285	[0.16, 0.05]	−0.09	.165	[−0.21, 0.04]	−0.03	.530	[−0.11, 0.06]	0.02	.748	[−0.11, 0.15]
English	0.06	.185	[0.03, 0.16]	0.10	.065	[−0.01, 0.21]	−0.03	.560	[−0.12, 0.06]	0.02	.780	[−0.10, 0.13]
Effort												
Math	0.13	.054	[0.00, 0.27]	0.07	.293	[−0.06, 0.21]	0.01	.830	[−0.11, 0.13]	0.04	.561	[−0.08, 0.16]
German	0.00	.978	[−0.09, 0.09]	−0.11	.109	[−0.24, 0.02]	−0.03	.525	[−0.13, 0.07]	−0.07	.366	[−0.21, 0.08]
English	0.13	.009	[0.03, 0.24]	0.07	.285	[−0.06, 0.21]	0.03	.604	[−0.08, 0.13]	0.08	.240	[−0.05, 0.21]

Note. Regression coefficients represent effect sizes at the classroom level that are standardized according to the total variance of the outcome variable. The respective constructs for math, German, and English at the pretest were included as covariates on the student level as well as the classroom level. CI = confidence interval.

with students in classes in the control condition at the posttest, controlling for their pretest self-concept in math, German, and English, β = .10, p = .049, 95% CI [0.00, 0.20]. The quotations condition did not show a significant effect on German (p = .285) or English (p = .185) self-concept at the posttest. No significant effects of the quotations condition on self-concept were found at the follow-up ($p \geq .065$). In line with the results for value, the text condition did not show any significant effect on students' domain-specific self-concepts at the posttest or the follow-up (all p's ≥ .446).

For effort, students in the quotations condition did not report higher effort in math than that of students in the control condition at the posttest (p = .054). In line with this, no effects of the quotations condition were observed for German effort (p = .978) at the posttest. However, at the posttest, students in the quotations condition reported higher effort in English versus students in the control condition, β = .13, p = .009, 95% CI [0.04, 0.24]. At the follow-up, no effects of the quotations condition on students' effort were observed (all p's ≥ .109). Again, we found no effect of the text condition on effort in any subject, neither at the posttest (all p's ≥ .525) nor at the follow-up (all p's ≥ .240).

We also examined whether the intervention effects varied depending on student gender or initial motivation. As our model

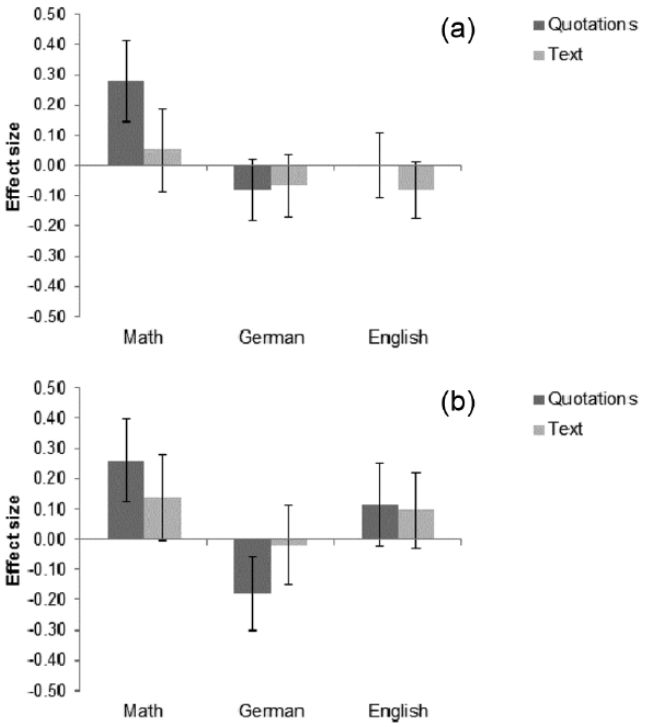

FIGURE 2. *Effects of the intervention conditions (as compared with the control condition) on value at (a) the posttest and (b) the follow-up. Error bars represent 95% confidence intervals.*

was already rather complex (i.e., multilevel analyses including multiple constructs), we could not investigate moderation effects using this same model. Instead, we used univariate analyses examining one subject at a time, with the respective pretest construct being used as a covariate, and we did not include any auxiliary variables. In line with the results reported by Gaspard, Dicke, Flunger, Brisson, et al. (2015), we found that the intervention conditions showed larger effects on math motivation for females than for males. Specifically, we found significant interaction effects between the text condition and gender on math value and self-concept at both the posttest and the follow-up and one significant interaction between the quotations condition and gender on math self-concept at the posttest. However, side effects in German and English did not vary by student gender. When examining whether intervention effects varied according to the initial level of the respective construct in math, German, or English, we only found one significant interaction effect (i.e., the effect of the text condition on English self-concept varied according to initial self-concept in English) for a total of 108 interaction effects being tested. The main effects of the intervention, however, were robust in these analyses.

Discussion

Within a cluster-randomized study aiming at fostering motivation in math, we addressed the important question of how motivational interventions in one subject affect motivational patterns across subjects. Based on the literature on dimensional comparisons, side effects were assessed in the verbal domain. The intervention condition that was successful in fostering math value showed negative effects on value for students' native language (i.e., German) 5 months after the intervention. No effects were found on value in English. Whereas this effect pattern was observed for value as the focal construct of the intervention, the effects did not generalize to students' self-concepts and effort.

Side Effects of Motivational Interventions

Whereas we found only one significant negative effect on motivation in the verbal domain (i.e., German value at the follow-up), the pattern of findings over conditions, constructs, and subjects was generally in line with our expectations. First, this side effect was found for the quotations condition, which was also effective in fostering math value. The text condition, however, did not significantly promote students' value beliefs in math. This is partly in line with the results reported before (Gaspard, Dicke, Flunger, Brisson, et al., 2015): Whereas the text condition did show effects on utility value, no effects were found for the other value components. In this article, we looked only at a composite value measure and did not find significant effects of the text condition. It is still not entirely clear which intervention strategies are most effective in promoting value beliefs, and this probably also depends on the characteristics of the specific sample (Durik, Hulleman, et al., 2015). In this specific case, it seems that evaluating quotations from potential role models was a task that fitted better to the needs and preferences of ninth-grade students than writing an essay about the relevance of a subject. With respect to side effects on German motivation, it seems worthwhile to consider the nature of the relevance-inducing tasks that students completed. Specifically, writing an essay is a task that is typically done in language arts classes. One might therefore think that such a task yields stronger effects for students who value language arts in general. Our results, however, did not support this hypothesis.

Second, we found positive main effects as well as negative side effects for the focal construct of the intervention (i.e., value). We did not, however, find the same effects for self-concept and effort. We found a positive effect of the quotations condition on only math self-concept at the posttest, which, however, did not sustain until the follow-up. Whereas effects of this short intervention on students' value beliefs were found 5 months after the intervention, it is possible that the effects of the intervention on students' value beliefs would need to be stronger to also find long-term effects on students' self-concept and effort in these subjects, and this would probably also increase the likelihood to find side effects for self-concept and effort. Against our expectations, we found a positive side effect of the quotations

condition on effort in English at the posttest, which could, however, no longer be found at the follow-up. One noticeable finding, which might contribute to this effect, is that students' reported effort seemed to be less domain specific, as it showed positive correlations across the three subjects (for similar results, see Trautwein, Lüdtke, Schnyder, & Niggli, 2006). This varying domain specificity of different motivational variables warrants future investigation.

Third, we examined side effects for two verbal subjects and found negative side effects for only one of them (i.e., German). In the literature, most support has been found for dimensional comparisons between math and students' native language (Möller & Marsh, 2013). Students in our study generally reported high value for English. The intervention aimed at triggering reflections on the relevance of the current learning material for future careers. In Germany, English is generally perceived as highly relevant for almost every career, including math-intensive fields, and students' knowledge about this might have buffered negative effects on English value. By comparison, the curriculum in German in secondary school is highly focused on literature, and students might therefore perceive its relevance as more limited (for students' perceptions of domain characteristics, see Goetz et al., 2014).

Fourth, when the negative effect of the quotations condition on German value as compared with the control condition was interpreted, the development over the course of the school year yields valuable information. Mean levels of German value observed in the different groups suggest that German value in the control condition increased from pretest to follow-up. This development might be affected by a number of factors inherent in the school context (e.g., timing of holidays, examinations, and content variations) as well as by factors pertaining to the study context that tapped all conditions (i.e., repeated questionnaire distribution within mathematics lessons). German value as reported in the quotations condition thus showed only a negative development as compared with this "natural" development. These negative effects on German value were found only at the follow-up, whereas effects on math were already found at the posttest. Dimensional comparison processes thus seem to rely on changes in motivation in the targeted domain occurring prior to spillover effects on other domains.

Theoretical Implications

Intervention studies that experimentally manipulate motivation in one domain and assess effects on nontargeted domains are one way to examine the role of dimensional comparisons in the school context. In addition to self-concept, where dimensional comparison effects have been extensively supported (Marsh & Hau, 2004; Möller et al., 2009), our study examined side effects on task value, which might be affected by similar dimensional comparison processes (see also Nagy et al., 2006; Nagy et al., 2008). This is in line with the assumptions of expectancy-value theory discussing the role of individual hierarchies for the value attached to different choice options (Eccles, 2009). Our results further suggest that contrast effects between math and the verbal domain depend on the specific subject. It is possible that such contrast effects are stronger for language arts than for foreign languages. However, the nature of the experimental manipulation, which focused on stimulating relevance reflections in this specific case, might also be important here.

Future research seems necessary to shed further light on the mechanisms that are behind side effects of motivational interventions. To do so, it would be necessary to collect more process-related measures focusing on dimensional comparisons. Qualitative analyses of students' responses to the intervention tasks might be one way to achieve such data (cf., Hulleman & Cordray, 2009). However, as students were explicitly instructed to focus on the relevance of math, these responses seem not to be helpful in the present study. Introspective studies using diary methods have found that students engage in dimensional comparisons in everyday life (Möller & Husemann, 2006), and similar approaches could also be used within intervention research.

Educational Implications

Our results have implications for education in general as well as for educational interventions in particular. Regarding the practical implications, the side effects found in our study would be judged as small when we apply conventional standards (Cohen, 1988). However, the intervention consisted of a 90-minute session in math classrooms and two short booster tasks. Compared with the cumulative effects of regular classroom experiences, this is a minimal intervention. Most notably, however, the intervention focused only on math. From this perspective, any side effect being found on motivation in nontargeted subjects can therefore be judged as quite meaningful. When transferring these results to the regular classroom setting, it seems also reasonable to assume that classroom experiences in one subject might affect student motivation in another subject (see Dietrich, Dicke, Kracke, & Noack, 2015, for first evidence of such effects with respect to teacher behavior). If students perceive one teacher as very good in connecting the learning material to their lives, this might have not only a positive impact on the value that they attach to this particular subject (see Wang, 2012) but also a negative impact on the value that they attach to another subject, where making connections is stimulated less frequently through the teacher.

Side effects as assessed in our study also have implications for further development of motivational interventions. At first sight, adverse effects of motivational interventions are troubling. However, if the ultimate goal of STEM interventions is to engage students in STEM, such effects might

be a risk that not only needs to be accepted but can actually help foster STEM engagement and career choices. As positive effects on math motivation as well as negative effects on verbal motivation positively affect intraindividual comparisons between math and verbal domains, these side effects could increase the likelihood for students to pursue math-related careers (cf. Parker et al., 2012). Future research using longer follow-up designs should assess such potential effects on later career choices.

Ethical considerations, however, seem to speak against diminishing motivation in one subject as a means to foster motivation for another subject. Given the role of students' value beliefs for student engagement and choices in the respective domains (Wigfield et al., 2009), value interventions in one particular subject can also be understood as pushing students into one direction or pulling them away from another. Researchers developing motivational interventions should therefore carefully consider side effects and seek for possibilities to foster motivation and engagement across subjects. One way to achieve this could be by developing broader interventions that stimulate the perceived relevance of various school subjects or school in general (for a similar intervention strategy, see Woolley, Rose, Orthner, Akos, & Jones-Sanpei, 2013). Alternatively, one could try to implement an inoculation strategy against side effects. As contrast effects seem to depend on students' beliefs about the association between mathematical and verbal abilities (Möller et al., 2006), targeting these beliefs would be one possible strategy. To do so, one way to go might be to present research results about the associations typically being found between mathematical and verbal achievement, on one hand, and mathematics and verbal self-concepts, on the other.

Limitations

Our study is the first to address side effects of motivational interventions on motivation in nontargeted domains. It raises a number of questions that future research should seek to address. When the results of our study are interpreted, the following limitations should be kept in mind. First, whereas we used parallel scales across subjects to directly compare the results, we had only short self-report scales available for the verbal subjects. This seems to be especially disadvantageous for the value scale. In a previous study using the same data (Gaspard, Dicke, Flunger, Brisson, et al., 2015), intervention effects were differentiated with stronger effects for math utility value than for attainment and intrinsic value. We were, however, not able to assess the same differential effects for German and English value. Also, we did not have any data on students' achievement or actual behavior in those subjects. Future research would need to take measures of students' behavior, achievement, and long-term choices into account to assess whether changes in value beliefs also result in behavioral change.

Second, we included two verbal domains that we thought of as the most important candidates for side effects and found negative effects for students' native language. To further examine which domains are affected through spillover effects and how, future research including a broader range of domains is needed. Depending on the similarity of domains, positive side effects via assimilation processes are also possible (Möller & Marsh, 2013). For the case of motivational interventions in the STEM domain, assimilation effects are plausible between math and sciences where such effects for students' self-concept have been reported in path-analytic studies (Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2015). However, the existence of such positive side effects might again depend on the content of the intervention and how students perceive these different subjects.

Third, our study is the first to address side effects of motivational interventions. Whereas we found some evidence of such side effects that was in line with our expectations, these side effects were limited to German value at the follow-up. This calls for replication in other intervention studies with adequate power to detect such effects. Although we tested our hypotheses in a large-scale intervention study, the sample size for this study was determined according to a power analysis for testing the main intervention effects. To examine side effects in intervention research, which would be expected to be smaller than the main intervention effects (cf., Möller et al., 2009), an even larger sample might be needed to achieve an adequate power. Although intervention studies in the field bring many advantages with them, interventions in the field often show smaller effects than interventions in the laboratory due to variations in the implementation and the context (Hulleman & Cordray, 2009) and therefore require large sample sizes to be able to achieve realistic estimates of effects (cf. Gelman & Carlin, 2014).

Appendix

Items Used to Assess Value, Self-Concept, and Effort

Value

1. It is important to me to be good at math/German/English.
2. It is important to me to know a lot of math/German/English.
3. I like doing math/German/English.
4. Math/German/English is useful to me.

Self-concept

1. I just have no talent for math/German/English. (reverse scored)
2. I am good at math/German/English.
3. Math/German/English just isn't my thing. (reverse scored)
4. Math/German/English comes naturally to me.

Effort

1. I really work hard in math/German/English.
2. I am very industrious in math/German/English.
3. I give my best in math/German/English.
4. I work very thoroughly on all math/German/English assignments and homework.

Acknowledgments

This research was funded in part by the German Research Foundation (TR 553/7-1) awarded to Ulrich Trautwein, Oliver Lüdtke, and Benjamin Nagengast. Brigitte Maria Brisson, Hanna Gaspard, and Isabelle Häfner were members of the Cooperative Research Training Group of the University of Education, Ludwigsburg, and the University of Tübingen, which was supported by the Ministry of Science, Research and the Arts in Baden-Württemberg. Hanna Gaspard and Isabelle Häfner were also doctoral students of the LEAD Graduate School (GSC 1028), funded by the Excellence Initiative of the German federal and state governments. We also acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen. We thank Katharina Allgaier and Evelin Herbein for their help conducting this research.

Note

1. We additionally conducted analyses assessing intervention effects on all outcomes using univariate models that included only the pretest value in the respective subject as a covariate instead of all three subjects. The pattern of results was very similar to that reported here: There was no substantial difference in the size of the intervention effects found, and the significance of the effects did not change.

References

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, *93*, 23–34. http://doi.org/10.I037//0022-0663.93.1.23

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

Brisson, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Nagengast, B., & Trautwein, U. (2016). *Short intervention, sustained effects: Promoting students' math competence beliefs, effort, and achievement*. Manuscript submitted for publication.

Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology*, *48*, 1612–1628. http://doi.org/10.1037/a0030194

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Dietrich, J., Dicke, A.-L., Kracke, B., & Noack, P. (2015). Teacher support and its influence on students' intrinsic value and effort: Dimensional comparison effects across subjects. *Learning and Instruction*, *39*, 45–54. http://doi.org/10.1016/j.learninstruc.2015.05.007

Durik, A. M., Hulleman, C. S., & Harackiewicz, J. M. (2015). One size fits some: Instructional enhancements to promote interest. In K. A. Renninger, M. Nieswandt, & 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. http://doi.org/10.1007/s11031-014-9419-0

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*, 78–89. http://doi.org/10.1080/00461520902832368

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* (pp. 74–146). San Francisco, CA: Freeman.

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

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

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

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

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651. http://doi.org/10.1177/1745691614551642

Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, *33*, 9–33. http://doi.org/10.1016/j.cedpsych.2006.12.002

Goetz, T., Haag, L., Lipnevich, A. A., Keller, M. M., Frenzel, A. C., & Collier, A. P. M. (2014). Between-domain relations of students' academic emotions and their judgments of school domain similarity. *Frontiers in Psychology*, *5*, 1153. http://doi.org/10.3389/fpsyg.2014.01153

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, *10*, 80–100. http://doi.org/10.1207/S15328007SEM1001

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

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*, 899–906. http://doi.org/10.1177/0956797611435530

Harackiewicz, J. M., Tibbetts, Y., Canning, E., & Hyde, J. S. (2014). Harnessing values to promote motivation in education. In S. A. Karabenick & T. Urdan (Eds.), *Advances in motivation and achievement: Vol. 18. Motivational interventions* (pp. 71–105). Bingley, England: Emerald Group.

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*, 88–110. http://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. http://doi.org/10.1037/a0019506

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

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology*, *107*, 1086–1101. http://doi.org/10.1037/edu0000021

Karabenick, S. A., & Urdan, T. (Eds.). (2014). *Advances in motivation and achievement: Vol. 18. Motivational interventions*. Bingley, England: Emerald Group.

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

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology*, *96*, 56–67. http://doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. S., Abdelfattah, F., & Jansen, M. (2015). Dimensional comparison theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction*, *35*, 16–32. http://doi.org/10.1016/j.learninstruc.2014.08.005

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & 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*, 764–802. http://doi.org/10.1080/00273170903333665

Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, *10*, 259–284. http://doi.org/10.1037/1082-989X.10.3.259

Meyer, W.-U. (1982). Indirect communications about perceived ability estimates. *Journal of Educational Psychology*, *74*, 888–897. http://doi.org/10.1037/0022-0663.74.6.888

Moher, D., Schulz, K. F., & Altman, D. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *BMC Medical Research Methodology*, *1*, 2. http://doi.org/10.1186/1471-2288-1-2

Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology*, *98*, 342–353. http://doi.org/10.1037/0022-0663.98.2.342

Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology*, *93*, 826–835. http://doi.org/10.1037//O022-O663.93.4.826

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*, 544–560. http://doi.org/10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*, 1129–1167. http://doi.org/10.3102/0034654309337522

Möller, J., Streblow, L., & Pohlmann, B. (2006). The belief in a negative interdependence of math and verbal abilities as determinant of academic self-concepts. *British Journal of Educational Psychology*, *76*, 57–70. http://doi.org/10.1348/000709905X37451

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide*. 7th ed. Los Angeles, CA: Muthén & Muthén.

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). Washington, DC: American Psychological Association.

Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, *12*, 323–345. http://doi.org/10.1080/13803610600765687

National Science Board. (2007). *A national action plan for addressing the critical needs of the U.S. science, technology, engineering, and mathematics education system*. Arlington, VA: National Science Foundation.

Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. *Developmental Psychology*, *48*, 1629–1642. http://doi.org/10.1037/a0029167

Pohlmann, B., & Möller, J. (2009). On the benefit of dimensional comparisons. *Journal of Educational Psychology*, *101*, 248–258. http://doi.org/10.1037/a0013151

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

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

Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, *34*, 32–41. http://doi.org/10.1016/j.learninstruc.2014.08.001

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

Steenkamp, J. E. M., & Baumgartner, H. (2009). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*, 78–90. http://doi.org/10.1086/209528

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*, 432–444. http://doi.org/10.1037/0022-0663.99.2.432

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology*, *90*, 334–349. http://doi.org/10.1037/0022-3514.90.2.334

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*, 1115–1128. http://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*, 438–456. http://doi.org/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, & F. Guay (Eds.), *Theory driving research: New wave perspectives on self-processes and human development* (pp. 233–249). Charlotte, NC: Information Age.

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, England: National Foundation for Educational Research.

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*, 1643–1657. http://doi.org/10.1037/a0027247

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*, 304–340. http://doi.org/10.1016/j.dr.2013.08.001

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.

Wigfield, A., Tonks, S., & Klauda, S. T. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–75). New York, NY: 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*, 1309–1335. http://doi.org/10.3102/0002831213488818

Authors

HANNA GASPARD is a postdoctoral research fellow at the Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastr. 6, 72072 Tübingen, Germany; hanna.gaspard@uni-tuebingen.de. She is a postdoctoral fellow in the Pathways to Adulthood Programme and supported by a Postdoctoral Fellowship for Leading Early Career Researchers by the Baden-Württemberg Stiftung. Her research focuses on the development of student motivation in regular classrooms and targeted interventions to foster student

ANNA-LENA DICKE is a postdoctoral research fellow at the School of Education at the University of California, Irvine. Her research focuses on understanding the influence of the educational context on students' motivation, interests and their career pathways.

BARBARA FLUNGER was a postdoctoral research fellow at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen and is now assistant professor of education at the University of Utrecht. Her research focuses on inter-individual differences in students' motivation and their association with students' academic outcomes.

ISABELLE HÄFNER is a postdoctoral research fellow at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. Her research focuses on parental influences on student motivation and the effectiveness of interventions to promote student motivation.

BRIGITTE M. BRISSON was a doctoral student at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen and is now a research assistant at the German Institute for International Educational Research in Frankfurt am Main. She is primarily interested in motivation research and in the implementation and evaluation of classroom-based motivation interventions.

ULRICH TRAUTWEIN is professor at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. His research interests include the development of student motivation, personality, academic effort, and achievement.

BENJAMIN NAGENGAST is professor at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. His research interests include quantitative methods (causal inference, latent variable models, multilevel modeling), educational effectiveness, the evaluation of educational interventions and motivation and academic self-concept.