Online Supplemental Material

The Value of Valuing Math: Longitudinal Links Between Students' Intrinsic, Attainment and Utility Values in Math and Math Grades

by Weidinger, Spinath, & Steinmayr, submitted, *Motivation Science*

Overview

Supplement 1. Description of the longitudinal project on the determinants of students' academic achievement at the end of secondary school

The 3-year longitudinal project on determinants of students' academic achievement at the end of secondary school (e.g., Bergold & Steinmayr, 2018) examined how cognitive and non-cognitive student characteristics relate to students' academic achievement (i.e., grades) from Grades 11 to 13 in German academic track schools ("Gymnasium"). The project was undertaken for the Habilitation ("second dissertation", one possible requirement for getting a full professorship in Germany) of one of the authors (Steinmayr, 2010). The aim was to collect data of 450 to 500 students. Therefore, five schools were recruited for participation in the project. Within each school, all 11th grade classes participated. The final sample comprised 476 students (232 boys and 244 girls) from two schools in North-Rhine Westphalia and three schools in Baden-Württemberg. The project started at the beginning of Grade 11 ($t_{11.1}$, 2008) when students were on average 16.43 (SD = 0.55) years old. The second measurement occasion took place at the beginning of the second term in Grade 11 ($t_{11.2}$, 2009). The last two measurement occasions followed at an interval of 1 year at the beginning of the second term in Grade 12 ($t_{12.2}$, 2010) and Grade 13 ($t_{13.2}$, 2011).

At $t_{11.1}$, students completed tests on their cognitive abilities, they answered questions on their age, gender, and social background, and they filled in a questionnaire on school-related behaviors at home (e.g., value of school, use of learning programs, talking about school with parents, number of books at home, private tutoring). Moreover, students' parents rated their child's cognitive abilities, knowledge, motivation, and personality. From $t_{11.1}$ to $t_{13.2}$, students filled in a questionnaire on their ability self-concepts and their intrinsic, attainment and utility values in school, math, German, physics, and chemistry, on their hope for success and fear of failure, on their goal orientations in school, their subjective well-being and need for cognition. Schools provided students' report card grades from Grades 10 to 13 and information on whether students' attended advanced or basic courses in the domains of math, German, physics, and chemistry in Grades 12 and 13. In this study, we focus on students' task values and grades in the domain of math because math is highly relevant for professional success in contemporary society. For example, more jobs require math skills than ever before (NMAP, 2008), and degrees in math intensive fields are associated with higher status careers (Ma & Johnson, 2008).

Supplement 2. Missing Value Analysis

As common for longitudinal studies, we found two kinds of missing data (see Little, 2013). First, missing data resulted when students missed one or more measurement occasions. The main reason for this "attrition" was illness. Besides this, School 5 dropped out of the project after $t_{11.2}$ for reasons not related to the investigation. Thus, for students from School 5, values and grades in math are only available for Grade 11 ($t_{10.2}$, $t_{11.1}$, and $t_{11.2}$). There were no significant differences in values and grades in math at $t_{10.2}$, $t_{11.1}$, and $t_{11.2}$ between students from School 5 and students from the other schools (V = 0.023, $F_{(9, 350)} = 0.922$, $p = .506$). Moreover, for students from School 4, math grades at the end of Grade 13 ($t_{13.2}$) were not available also for reasons not related to the investigation. There were no significant differences in values and grades in math between students from School 4 and the other students in the sample (V = 0.107, $F_{(18, 168)} = 1.119$, $p = .338$). More detailed information about the number of students participating at each measurement point can be found in Table 1. Second, missing data resulted from nonresponse that occurred at each measurement occasion and ranged from 0.0% to 2.1% for all value components ($t_{11.1}$: 0.0%; $t_{11.2}$: 2.1%; $t_{12.2}$: 0.3%; $t_{13.2}$: 0.7%), from 0.0% to 12.2% for math grades ($t_{10.2}$: 0.0%; $t_{11.1}$: 7.4%; $t_{11.2}$: 9.5%. $t_{12.1}$: 10.0%; $t_{12.2}$: 10.0%; $t_{13.1}$: 10.3%; $t_{13.2}$: 12.2%), and from 1.0% to 6.7% for the covariates (cognitive abilities: 1.0%; math course selection: 6.7%; SES: 2.4%). Because full information maximum likelihood estimation (FIML) improves the accuracy and the power of the analyses in comparison to other methods to handle missing data and because it is regarded as state-of-the-art missing data technique (Graham & Coffman, 2012; Little, 2013; Schafer & Graham, 2002), we used this model-based approach to handle missing data in our analyses.

Supplement 3. Testing longitudinal measurement invariance of the value components

To test for longitudinal measurement invariance of the three value components, we estimated a longitudinal confirmatory factor analysis (CFA) model for each value component. In this model the value component was indicated by three items at each measurement occasion, the residual variances among the corresponding indicators were allowed to correlate over time, and the effects coding method of identification was used for scale setting (see Little, 2013). We used a stepwise approach and first compared a model showing configural invariance (i.e., only the factor structure was invariant across time) with a more parsimonious model showing weak invariance (i.e., the factor loadings of the manifest indicators were fixed to invariance across time). If the change in model fit was negligible, in a second step, we additionally fixed the intercepts of the manifest indicators to be invariant across time (strong invariance). If this additional restriction did not lead to a significant deterioration in model fit, we concluded that the supposition of strong factorial invariance holds. To evaluate the invariance constraints, we calculated the chi-square difference test ($\Delta\chi^2$, Δdf) and the change in CFI (ΔCFI; see Little, 2013). According to findings from a simulation study, the supposition of invariance is tenable if $\Delta CFI \leq .01$ (Cheung & Rensvold, 2002). Constraining the factor loadings and the intercepts of the manifest indicators of intrinsic, attainment and utility values to be invariant over time, did not lead to substantial declines in model fit (see Table S3).

Table S3

Model Fit Statistics for Testing the Longitudinal Measurement Invariance of Intrinsic, Attainment, and Utility Values in Math

Model tested	χ^2 (*df*)	*p*	$\Delta\chi^2$ (*df*)	*p*	RMSEA	RMSEA 90% CI	SRMR	CFI	ΔCFI	Pass?
Intrinsic Value Math										
Configural invariance	38.845 (30)	.129	—	—	.026	[.000, .047]	.016	.998	—	—
Weak invariance	40.277 (36)	.290	1.432 (6)	.964	.017	[.000, .039]	.016	.990	.008	Yes
Strong invariance	40.340 (42)	.544	0.063 (6)	.999	.000	[.000, .031]	.016	> .999	.009	Yes
Attainment Value Math										
Configural invariance	45.097 (30)	.038	—	—	.034	[.008, .054]	.027	.996	—	—
Weak invariance	65.042 (36)	.002	19.945 (6)	.003	.043	[.026, .060]	.038	.992	.004	Yes
Strong invariance	65.117 (42)	.013	0.075 (6)	.999	.036	[.017, .052]	.038	.994	.002	Yes
Utility Value Math										
Configural invariance	71.184 (30)	< .001	—	—	.056	[.040, .073]	.037	.985	—	—
Weak invariance	77.076 (36)	< .001	5.892 (6)	.435	.051	[.036, .067]	.037	.986	.001	Yes
Strong invariance	77.109 (42)	.001	0.033 (6)	.999	.044	[.028, .059]	.037	.988	.002	Yes

Note. *N* = 429. Math values were assessed at four measurement occasions with three items. All variables were class wise *z*-standardized. Maximum likelihood estimation (ML) was used. χ^2(*df*) = chi-square test statistic with degrees of freedom in parentheses; $\Delta\chi^2$ (*df*) = test statistic of the chi-square difference test with degrees of freedom in parentheses; RMSEA = root mean square error of approximation along with its associated confidence interval; SRMR = standardized root mean square residual; CFI = comparative fit index; Configural invariance = invariant factor structure; Weak invariance = invariant factor loadings; Strong invariance = invariant factor loadings and intercepts.

Supplement 4. Unconstrained overall model

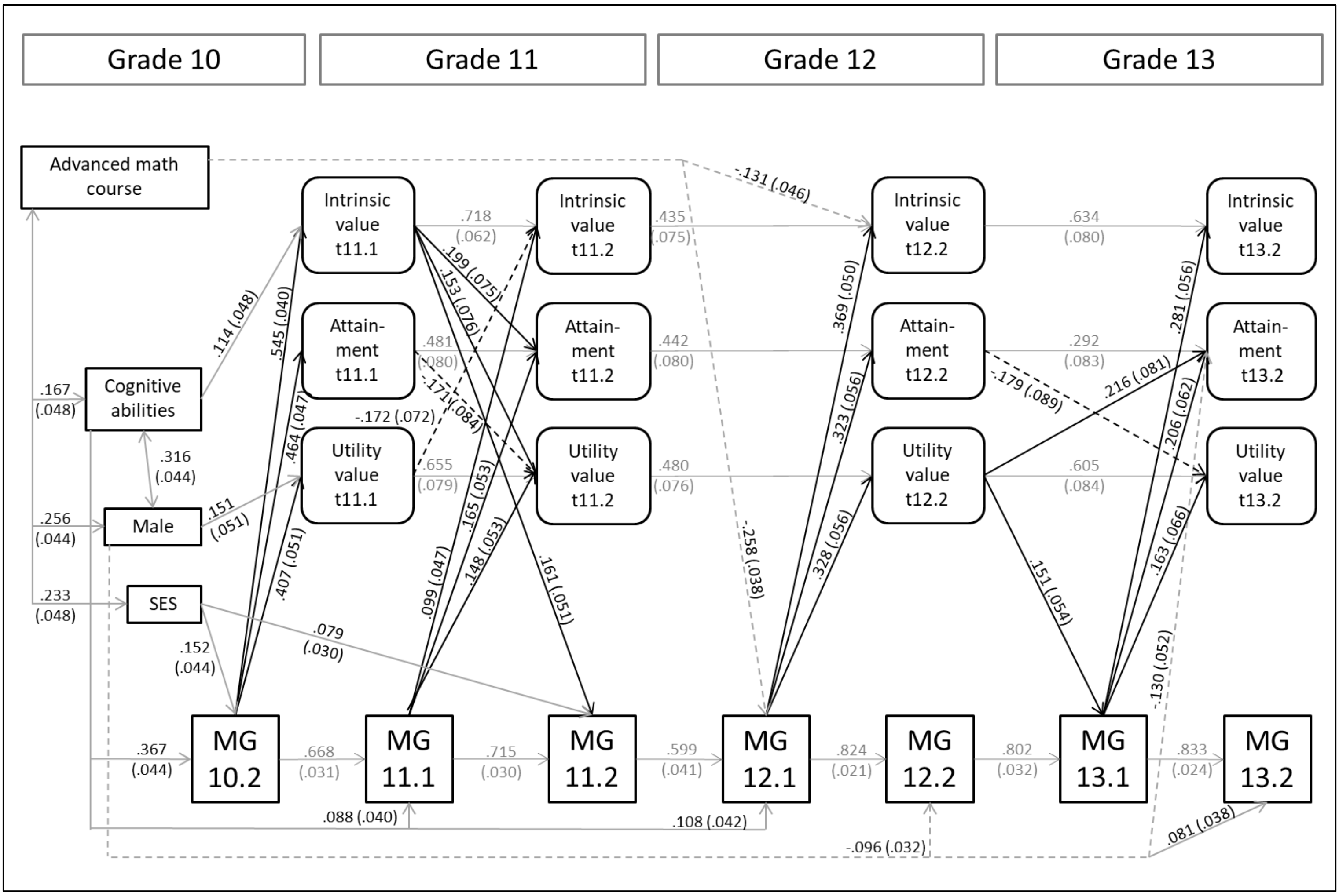

Figure S4. Unconstrained overall model with students' intrinsic, attainment, and utility values in math and math grades (MG) from Grade 10 to Grade 13. Students' gender (male = 1 vs. female = 0), socio-economic status (SES), cognitive abilities, and advanced math course selection (advanced course = 1 vs. basic course = 0) were included as covariates. $N = 476$. Standardized solution (standard errors in parentheses). Only significant paths are depicted ($p < .05$). Grey coefficients represent autoregressive relations; grey lines indicate autoregressive paths and paths leading from covariates to other variables; dashed lines indicate negative effects. All variables were class- or course-wise z-standardized. t = measurement occasion. Values were assessed at four measurement occasions with 3 items; factor loadings and intercepts of the manifest indicators were invariant across time, and residuals of all corresponding value items collected at subsequent measurement occasions were allowed to correlate. Measurement models, phantom constructs, and residual correlations between endogenous variables are not shown. Model fit: $\chi^2 = 1526.036$, $df = 876$, $p < .001$; RMSEA = .039 [.036, .043], CFI = .956; SRMR = .054.

Supplement 5. Single models for intrinsic, attainment and utility values

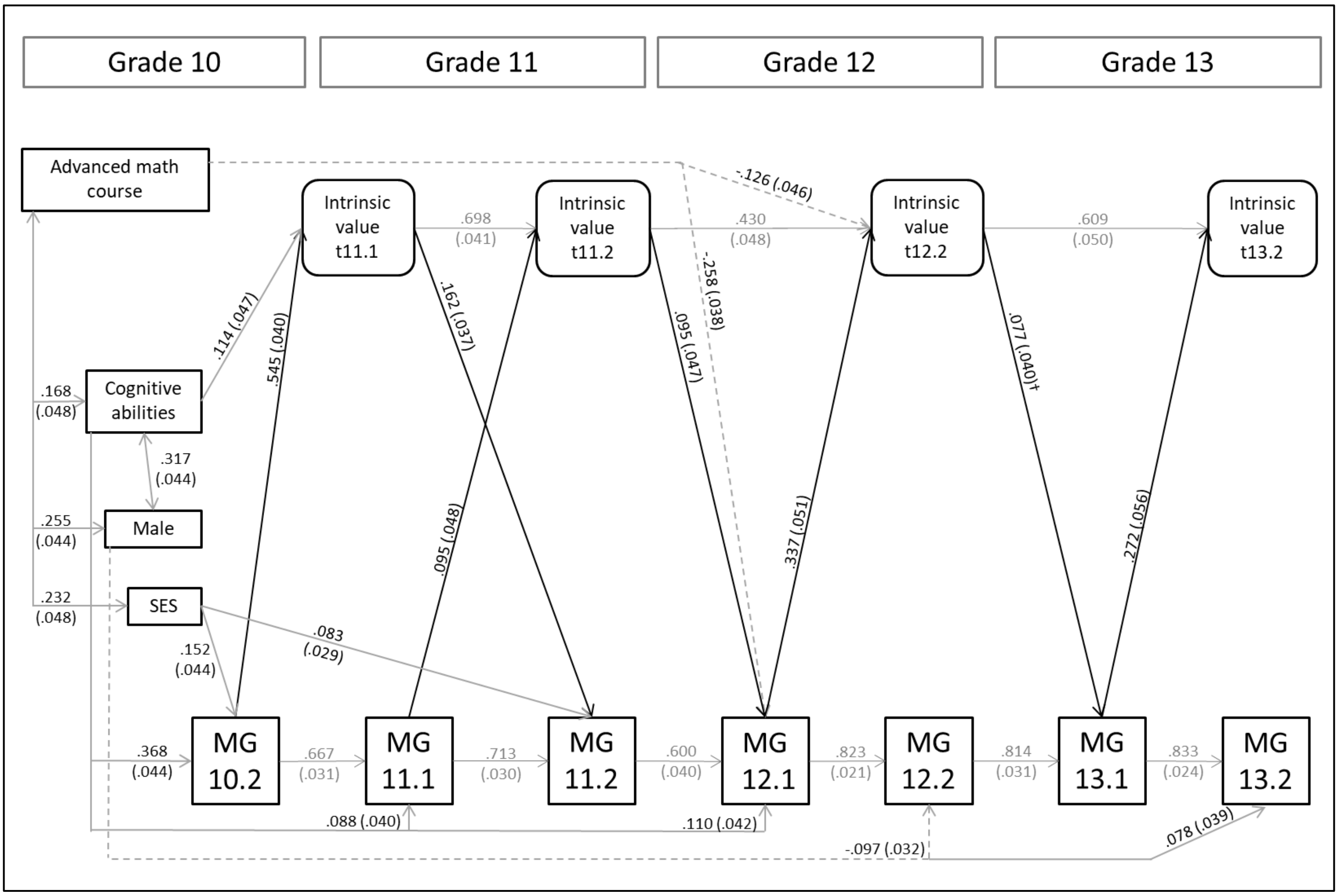

Figure S5.1. Single model with students' intrinsic value in math and math grades (MG) from Grade 10 to Grade 13. Students' gender (male = 1 vs. female = 0), socio-economic status (SES), cognitive abilities, and advanced math course selection (advanced course = 1 vs. basic course = 0) were included as covariates. $N = 476$. Standardized solution (standard errors in parentheses). Only significant paths are depicted ($p < .05$). Grey coefficients represent autoregressive relations; grey lines indicate autoregressive paths and paths leading from covariates to other variables; dashed lines indicate negative effects. All variables were class- or course-wise *z*-standardized. t = measurement occasion. Intrinsic value was assessed at four measurement occasions with 3 items; factor loadings and intercepts of the manifest indicators were invariant across time, and residuals of all corresponding value items collected at subsequent measurement occasions were allowed to correlate. Measurement models, phantom constructs, and residual correlations between endogenous variables are not shown. Model fit: $\chi^2 = 489.375$, $df = 170$, $p < .001$; RMSEA = .063 [.056, .069], CFI = .952; SRMR = .058. †$p = .056$.

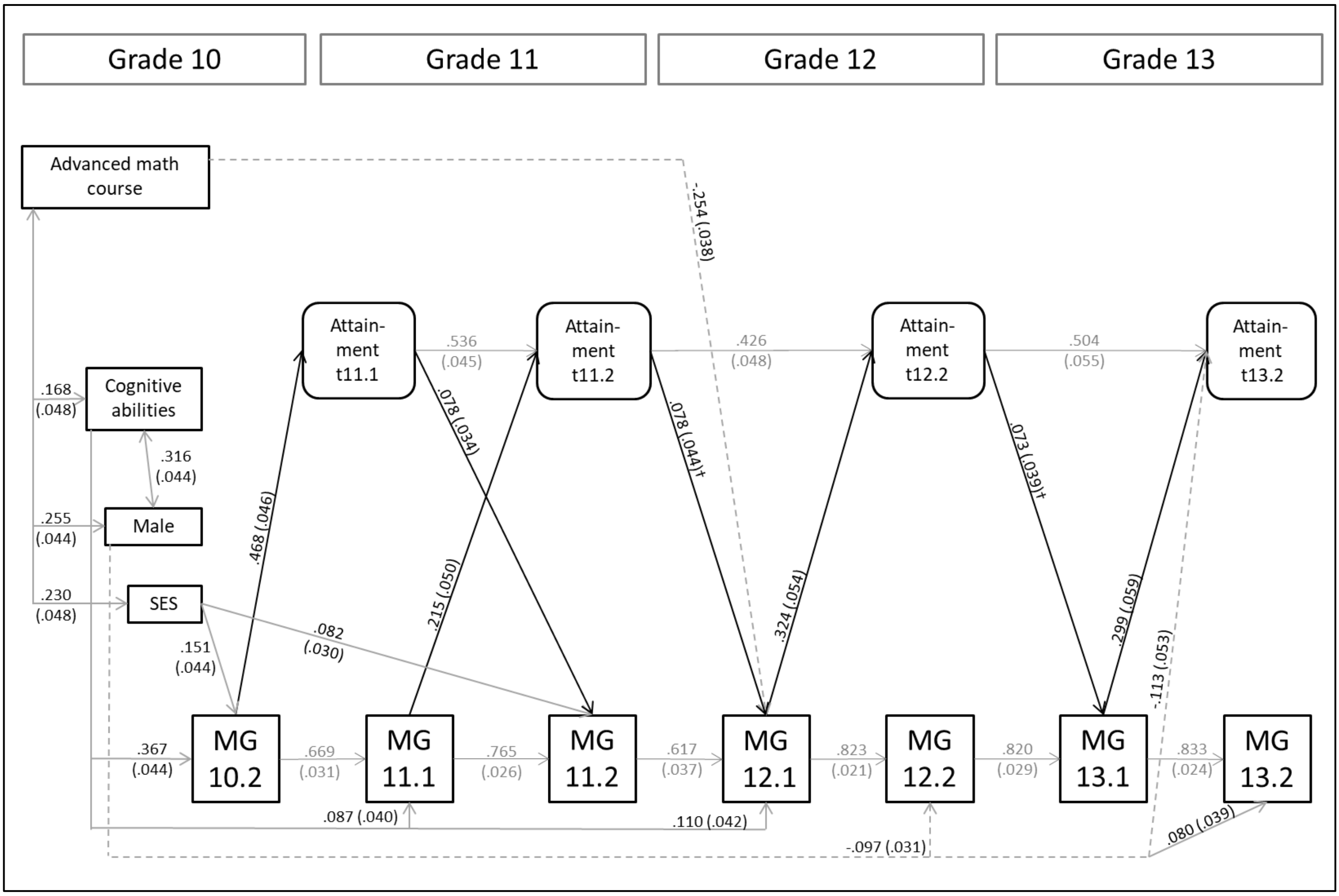

Figure S5.2. Single model with students' attainment value in math and math grades (MG) from Grade 10 to Grade 13. Students' gender (male = 1 vs. female = 0), socio-economic status (SES), cognitive abilities, and advanced math course selection (advanced course = 1 vs. basic course = 0) were included as covariates. N = 476. Standardized solution (standard errors in parentheses). Only significant paths are depicted ($p < .05$). Grey coefficients represent autoregressive relations; grey lines indicate autoregressive paths and paths leading from covariates to other variables; dashed lines indicate negative effects. All variables were class- or course-wise z-standardized. t = measurement occasion. Attainment value was assessed at four measurement occasions with 3 items; factor loadings and intercepts of the manifest indicators were invariant across time, and residuals of all corresponding value items collected at subsequent measurement occasions were allowed to correlate. Measurement models, phantom constructs, and residual correlations between endogenous variables are not shown. Model fit: χ^2 = 493.487, df = 170, $p < .001$; RMSEA = .063 [.057, .070], CFI = .948; SRMR = .064. †$p < .073$.

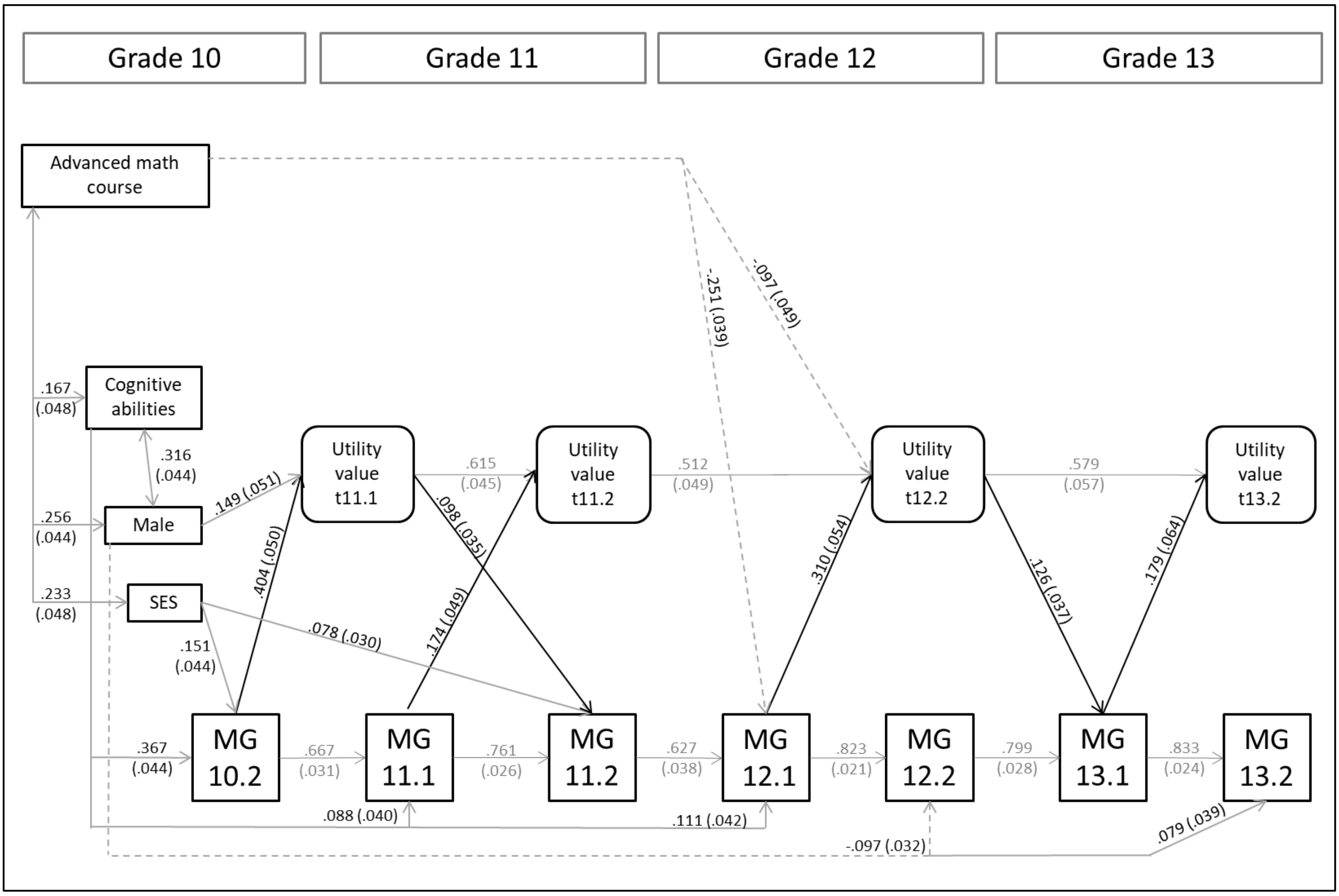

Figure S5.3. Single model with students' utility value in math and math grades (MG) from Grade 10 to Grade 13. Students' gender (male = 1 vs. female = 0), socio-economic status (SES), cognitive abilities, and advanced math course selection (advanced course = 1 vs. basic course = 0) were included as covariates. $N = 476$. Standardized solution (standard errors in parentheses). Only significant paths are depicted ($p < .05$). Grey coefficients represent autoregressive relations; grey lines indicate autoregressive paths and paths leading from covariates to other variables; dashed lines indicate negative effects. All variables were class- or course-wise *z*-standardized. t = measurement occasion. Utility value was assessed at four measurement occasions with 3 items; factor loadings and intercepts of the manifest indicators were invariant across time, and residuals of all corresponding value items collected at subsequent measurement occasions were allowed to correlate. Measurement models, phantom constructs, and residual correlations between endogenous variables are not shown. Model fit: $\chi^2 = 491.557$, $df = 170$, $p < .001$; RMSEA = .063 [.057, .070], CFI = .939; SRMR = .056.

Supplement 6. Constrained overall model for students in advanced math courses

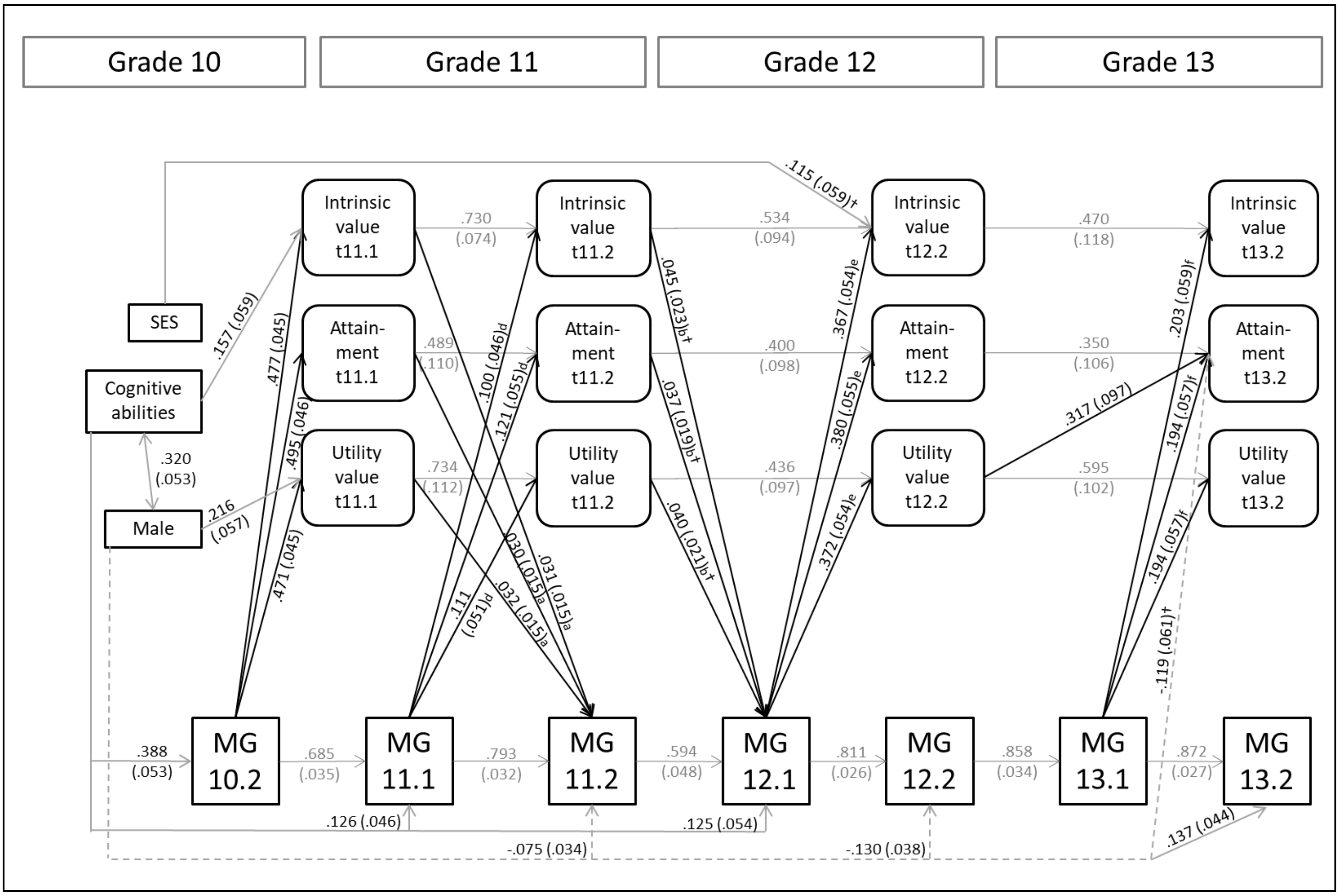

Figure S6. Constrained overall model for students in advanced math courses with students' intrinsic, attainment, and utility values in math and math grades (MG) from Grade 10 to Grade 13. Students' gender (male = 1 vs. female = 0), socio-economic status (SES), and cognitive abilities were included as covariates. $N = 310$. Standardized solution (standard errors in parentheses). Only significant paths are depicted ($p < .05$). Grey coefficients represent autoregressive relations; grey lines indicate autoregressive paths and paths leading from covariates to other variables; dashed lines indicate negative effects. All variables were class- or course-wise z-standardized. t = measurement occasion. Values were assessed at four measurement occasions with 3 items; factor loadings and intercepts of the manifest indicators were invariant across time, and residuals of all corresponding value items collected at subsequent measurement occasions were allowed to correlate. Measurement models, phantom constructs, and residual correlations between endogenous variables are not shown. Model fit: $\chi^2 = 1421.077$, $df = 857$, $p < .001$; RMSEA = .046 [.042, .050], CFI = .945; SRMR = .069. †p = .051.