

Appendix A: Correlations between reported experiences/beliefs (item/factor scores)

Item/factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1) Self-concept	1																
(2) Self-efficacy	.547	1															
(3) Mastery experiences (current grade)	.488	.626	1														
(4) Mastery norms (what is a good grade)	^{P5} .065	.309	.291	1													
(5) Subject-comparison	.495	.391	.321	^N .026	1												
(6) Peer-comparison	.494	.379	.314	^N .039	.625	1											
(7) Anxiety (absence of)	.548	.416	.360	^N .035	.666	.666	1										
(8) Praise (social persuasions)	.592	.401	.383	^N .045	.343	.359	.391	1									
(9) Vicarious experiences	.354	.300	.222	^{P5} .053	.198	.221	.235	.415	1								
(10) Interest value	.624	.471	.430	^{P1} .086	.441	.345	.452	.569	.381	1							
(11) Utility value	.491	.398	.339	.099	.298	.230	.250	.450	.341	.717	1						
(12) Personal value	.480	.374	.343	.097	.309	.268	.285	.483	.320	.655	.717	1					
(13) Cost value (absence of)	^N .012	^N -.020	^N -.051	^N -.035	.158	.141	.211	^N -.047	-.096	-.127	-.270	-.287	1				
(14) Teacher perceptions	.410	.235	.242	^N .048	.246	.213	.305	.507	.293	.606	.488	.423	^{P1} -.085	1			
(15) Gender (1=male)	.195	.289	.268	.149	.221	.171	.265	.152	.120	.205	.109	.200	^N .022	^{P1} .081	1		
(16) Task score	.341	.436	.507	.145	.265	.230	.306	.271	.169	.350	.240	.213	^N .017	.237	.191	1	
(17) Task confidence	.557	.526	.524	.166	.376	.356	.457	.411	.289	.530	.410	.417	^N -.023	.357	.351	.521	1
(18) Task confidence bias	.132	^N -.002	^{P5} -.063	^N -.009	^{P5} .061	^{P1} .080	^{P1} .085	^{P1} .085	^{P1} .084	.113	.120	.161	^N -.039	^{P1} .078	.102	-.634	.329

Notes: Correlations (Pearson R coefficients) were all significant at $p < .001$, except when highlighted with superscript indicators: *P1* significance at $p < .01$ only; *P5* significance at $p < .05$ only; *N* over $p < .05$ (not significant).

Appendix B: Mediators of the predictive association between students' reported grades and self-concept or self-efficacy beliefs

Conceptually, 'mediators' explain or account for the relation between a factor and an outcome: the factor predicts the mediator, which in turn predicts the outcome, thereby reducing the direct effect from the factor to the outcome (Baron & Kenny, 1986). Sobel tests (Sobel, 1982) have been established as reliably identifying when mediation occurs (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), and these confirmed that all the items/factors, with the exceptions of mastery norms and gender, individually mediated the predictive association between students' mastery experiences (current grades) and their self-concept or self-efficacy beliefs.

Mediator	Current grade → Self-concept			Current grade → Self-efficacy		
	Sobel sig. (p)	Proportion of total effect mediated	Ratio of indirect to direct effect	Sobel sig. (p)	Proportion of total effect mediated	Ratio of indirect to direct effect
Mastery norms (what is a good grade)	.018	-.010	-.010	.152	.022	.023
Subject-comparison	<.001	.229	.297	<.001	.149	.175
Peer-comparison	<.001	.221	.284	<.001	.142	.166
Anxiety (absence of)	<.001	.313	.455	<.001	.179	.217
Praise (social persuasions)	<.001	.362	.567	<.001	.180	.220
Vicarious experiences	.007	.110	.123	<.001	.069	.075
Interest value	<.001	.417	.714	<.001	.233	.304
Utility value	<.001	.239	.315	<.001	.167	.201
Personal value	<.001	.199	.249	<.001	.124	.142
Cost value (absence of)	<.001	<.001	<.001	<.001	<.001	<.001
Teacher perceptions	<.001	.149	.175	<.001	.058	.061
Gender (1=male)	.262	.021	.022	.043	.018	.019
Self-efficacy / self-concept	<.001	.432	.760	<.001	.398	.662

Notes: Sobel tests considered the various coefficients and standard errors within three predictive models in order to determine whether mediation significantly occurred: (1) grade → mediator; (2) mediator and grade → self-concept (or self-efficacy); (3) grade → self-concept (or self-efficacy).

References

- Baron, R., & Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. doi:10.1037/0022-3514.51.6.1173
- MacKinnon, D., Lockwood, C., Hoffman, J., West, S., & Sheets, V. (2002). A Comparison of Methods to Test Mediation and Other Intervening Variable Effects. *Psychological Methods*, 7(1), 83–104. doi:10.1037//1082-989X.7.1.83
- Sobel, M. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, 13, 290–312. doi:10.2307/270723

Appendix C: Science items/factors predicting students' science self-concept beliefs (with self-efficacy)

Item/factor	Step 1				Step 2				Step 3				Step 4			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Constant/intercept	1.10	.14	<.001	NA	1.53	.16	<.001	NA	.88	.16	<.001	NA	-.19	.18	.308	NA
Self-efficacy	.45	.03	<.001	.964	.47	.03	<.001	1.003	.35	.03	<.001	.756	.21	.03	<.001	.441
Mastery experiences (current grade)	.21	.02	<.001	.622	.21	.02	<.001	.621	.17	.02	<.001	.499	.12	.02	<.001	.355
Mastery norms (what is a good grade)					-.12	.03	<.001	-.218	-.07	.02	.004	-.134	-.05	.02	.024	-.096
Subject-comparison									.13	.02	<.001	.373	.03	.02	.083	.099
Peer-comparison									.14	.02	<.001	.350	.06	.02	.011	.142
Anxiety (absence of)													.10	.03	<.001	.232
Praise (social persuasions)													.20	.02	<.001	.432
Vicarious experiences													.02	.02	.382	.039
Interest value													.13	.03	<.001	.298
Utility value													.09	.03	.006	.188
Personal value													.01	.02	.598	.032
Cost value (absence of)													.02	.02	.232	.052
Teacher perceptions													.06	.03	.030	.112
Gender (1=male)													.07	.05	.175	.061
Explained variance	38.4%				39.9%				49.7%				64.4%			
Unexplained variance, school level	3.1%				2.3%				1.7%				1.8%			
Unexplained variance, residual	58.5%				57.8%				48.7%				33.8%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Items/factors used 1-6 scales unless otherwise indicated. Significant predictors ($p < .05$ or below) have been highlighted in bold for clarity. Unexplained variance at the residual level can be assumed to reflect the student level.

Appendix D: Science items/factors predicting students' science self-efficacy beliefs (with self-concept)

Item/factor	Step 1				Step 2				Step 3				Step 4			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Constant/intercept	2.31	.13	<.001	NA	1.66	.16	<.001	NA	1.42	.18	<.001	NA	1.44	.21	<.001	NA
Self-concept	.38	.02	<.001	.713	.39	.02	<.001	.728	.33	.03	<.001	.622	.27	.04	<.001	.512
Mastery experiences (current grade)	.21	.02	<.001	.582	.20	.02	<.001	.557	.17	.02	<.001	.482	.15	.02	<.001	.420
Mastery norms (what is a good grade)					.15	.02	<.001	.258	.14	.02	<.001	.238	.13	.03	<.001	.219
Subject-comparison													.04	.02	.106	.099
Peer-comparison													.07	.03	.012	.151
Anxiety (absence of)									.08	.02	<.001	.174	-.01	.03	.836	-.014
Praise (social persuasions)									.03	.03	.303	.054	.04	.03	.208	.074
Vicarious experiences									.04	.02	.039	.091	.01	.02	.494	.032
Interest value													.06	.04	.133	.119
Utility value													.12	.04	.001	.239
Personal value													-.02	.03	.462	-.047
Cost value (absence of)													.02	.02	.366	.042
Teacher perceptions													-.10	.03	.002	-.174
Gender (1=male)													.12	.06	.032	.104
Explained variance	45.8%				48.6%				50.2%				52.1%			
Unexplained variance, school level	5.5%				4.2%				4.4%				4.0%			
Unexplained variance, residual	48.7%				47.2%				45.4%				43.9%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Items/factors used 1-6 scales unless otherwise indicated. Significant predictors ($p < .05$ or below) have been highlighted in bold for clarity. Unexplained variance at the residual level can be assumed to reflect the student level.

Appendix E: Science items/factors predicting students' science self-concept beliefs across confidence bias groups (with self-efficacy)

Item/factor	Under-confident (U)				Accurate (A)				Over-confident (O)			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Constant/intercept	.24	.31	.446	NA	.14	.27	.600	NA	-.09	.31	.775	NA
Self-efficacy	^{uo} .10	.05	.042	.217	.19	.04	<.001	.408	^{uo} .26	.05	<.001	.550
Mastery experiences (current grade)	.13	.04	<.001	.363	.13	.03	<.001	.404	.10	.03	.002	.306
Mastery norms (what is a good grade)	-.08	.04	.050	-.151	-.10	.04	.006	-.175	-.04	.04	.331	-.077
Subject-comparison	^{ua} .13	.04	.001	.369	^{ua} -.01	.03	.620	-.041	.04	.04	.345	.101
Peer-comparison	.09	.04	.015	.243	.07	.03	.038	.174	-.02	.05	.720	-.039
Anxiety (absence of)	.14	.05	.006	.312	.13	.04	.001	.303	.02	.05	.756	.036
Social persuasions	.18	.05	<.001	.386	.20	.03	<.001	.433	.16	.05	.001	.342
Vicarious experiences	.02	.04	.609	.044	.01	.03	.682	.025	.04	.04	.267	.100
Interest value	.12	.07	.089	.277	.14	.05	.002	.327	.16	.06	.005	.366
Utility value	.02	.06	.712	.050	.04	.04	.327	.095	.20	.06	.002	.406
Personal value	.05	.04	.214	.144	-.01	.03	.865	-.014	-.03	.04	.561	-.067
Cost value (absence of)	.00	.03	.963	.004	.05	.03	.056	.122	.02	.04	.618	.044
Teacher perceptions	.01	.05	.790	.026	.06	.04	.139	.111	.08	.05	.127	.154
Gender (1=male)	^{ua} -.14	.09	.149	-.128	^{ua} .09	.07	.205	.084	.03	.09	.763	.023
Explained variance	62.7%				61.9%				71.3%			
Unexplained variance, school level	1.2%				3.7%				.5%			
Unexplained variance, residual	36.1%				34.4%				28.3%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Items/factors used 1-6 scales unless otherwise indicated. Significant predictors ($p < .05$ or below) have been highlighted in bold for clarity. Unexplained variance at the residual level can be assumed to reflect the student level. Significant differences ($p < .05$ or below) in coefficient magnitudes across groups (from separate interaction/moderation models for the various pairs of groups) have been highlighted in superscript.

Appendix F: Science items/factors predicting students' science self-efficacy beliefs across confidence bias groups (with self-concept)

Item/factor	Under-confident (U)				Accurate (A)				Over-confident (O)			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Constant/intercept	1.24	.38	.001	NA	1.70	.30	<.001	NA	.89	.39	.025	NA
Self-concept	^{uo} .16	.07	.029	.286	.25	.05	<.001	.468	^{uo} .41	.07	<.001	.773
Mastery experiences (current grade)	.20	.04	<.001	.522	.19	.03	<.001	.557	.10	.04	.016	.285
Mastery norms (what is a good grade)	.09	.05	.083	.148	.10	.04	.016	.166	.19	.05	<.001	.360
Subject-comparison	^{ua} .14	.05	.003	.360	^{ua} .02	.03	.612	.045	.01	.05	.895	.016
Peer-comparison	.08	.05	.073	.198	.05	.04	.167	.125	.06	.06	.277	.136
Anxiety (absence of)	-.07	.06	.250	-.146	.00	.05	.996	<.001	.05	.06	.409	.108
Social persuasions	.00	.06	.951	-.007	.07	.04	.095	.141	.05	.06	.397	.104
Vicarious experiences	.06	.04	.184	.127	.04	.03	.233	.080	-.05	.05	.263	-.115
Interest value	.04	.08	.657	.080	.00	.05	.949	-.007	.09	.07	.217	.183
Utility value	.11	.07	.121	.231	.11	.05	.027	.230	.10	.08	.193	.193
Personal value	^{ua} ^{uo} .10	.05	.058	.242	^{ua} -.02	.04	.485	-.064	^{uo} -.07	.05	.178	-.180
Cost value (absence of)	.02	.04	.648	.039	.01	.03	.732	.023	.05	.04	.283	.107
Teacher perceptions	-.06	.07	.379	-.097	-.09	.05	.043	-.162	-.12	.06	.064	-.213
Gender (1=male)	.09	.12	.448	.076	.19	.08	.020	.166	.14	.11	.214	.110
Explained variance	50.8%				49.7%				59.2%			
Unexplained variance, school level	3.3%				3.6%				3.8%			
Unexplained variance, residual	45.9%				46.7%				37.0%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Items/factors used 1-6 scales unless otherwise indicated. Significant predictors ($p < .05$ or below) have been highlighted in bold for clarity. Unexplained variance at the residual level can be assumed to reflect the student level. Significant differences ($p < .05$ or below) in coefficient magnitudes across groups (from separate interaction/moderation models for the various pairs of groups) have been highlighted in superscript.