Predictors of Test Anxiety Under 15-Year Old Students: A Comparative Study Between the Netherlands and Australia

Master Thesis Social Policy and Public Health at Utrecht University
Tim de Vries (3920852)
Supervisor: Dr. Carlijn Kamphuis
Wordcount: 7 509

Abstract:

Introduction: Test anxiety [TA] is a big issue in the Netherlands and Australia for both adolescents and adults. More than 50 per cent of students say they are anxious about their study. TA is correlated with other psychological issues: lower ambition, less teacher engagement, more perceived bullying and less perceived sense of belonging in the classroom. Gender also plays a role; girls seem to be more test-anxious than boys. This study seeks the effect of predictors of TA and the correlation of country and TA.

Method: 19 915 fifteen-year-old students are asked questions about TA, gender, ambition, sense of belonging, teacher engagement and bullying in the PISA 2015 study. Statistical programme R is used.

Results: The mediation analysis shows that the predictors of TA play a small but significant role in mediating country and TA. The direct effect of being Australian on TA is positively correlated. Being female is also highly positively correlated with TA, but not with the mediating factors.

Discussion and implications: This study is the first trying to explain TA through contextual factors. However, it is still unknown which contextual factors (and other factors) play a role in the correlation between TA and country of the student. This study can be the beginning of an evaluation of existing literature of TA, and possible policymaking with contextual factors of the country in mind.

Keywords: Test anxiety; Netherlands; Australia; adolescence; anxiety

Test Anxiety – Master Thesis Tim de Vries

Up to 41 per cent of children experience anxiety problems worldwide, making it one of the most widespread mental health problems (Von der Embse, Jester, Roy and Pose, 2018). Australia has one of the highest prevalence rates of anxiety disorders and depression in the world. Around seven per cent of the Australian population suffers from an anxiety disorder (World Health Organisation [WHO], 2017; Lawrence et al., 2015). This holds for both adults and children. Lawrence et al. (2015) found that 6.9% of Australian children between 4 and 17 are diagnosed with an anxiety disorder. This is equivalent to 278,000 Australian youths. Almost twenty per cent of Dutch 15-year-olds, who are often cited as the happiest on the planet (Currie et al., 2012; UNICEF, 2013), report to have anxiety problems (Stevens et al., 2018).

According to the Programme for International Student Assessment (PISA) questionnaire (Organisation for Economic Cooperation and Development [OECD], 2015), one of the most common anxieties among 15-year old students is schoolwork-related anxieties. Fifty-five per cent of the students report that they experience exam stress, even from a well-prepared test (OECD, 2015). In Australia, those schoolwork-related anxiety problems are problematic. Australian children have the most prevalence of school-related anxiety in the world (OECD, 2015): more than 67 per cent of Australian students feel anxious before a well-prepared test (OECD, 2015). In some other countries, however, school-related anxiety is less common. For example, Dutch children have the lowest school-related anxiety rates of Europe (Currie et al., 2012) and the world (OECD, 2015) However, even in the Netherlands, school stress is common under students (Stevens et al., 2018).

In the literature, school-related anxiety has multiple definitions. For example, Putwain, Chamberlain, Daly and Sadreddini (2014) describe it as a situation-specific form of trait anxiety: the perception of potential academic failure is threatening the self-worth of students, and they become anxious. Von der Embse et al. (2018) describes test anxiety (TA) as stress for exams or tests that: "*(…) [is] synonymous with the fear, worry or negative emotionality, which leads to negative behavioural, psychological or emotional responses*". Boehme et al. (2017) say that TA can be defined as behavioural, psychological and physiological reactions regarding worrying, related to the potential failure to achieve something. In this study, we use the definition of TA that the OECD uses in their Programme for International Student Assessment [PISA] 2015 research, which is: "*(…) [school-attending*

adolescents] are required to manage increasing academic demands in relatively more formal classroom settings. The pressure to get higher marks and the concern about receiving poor grades are some of the sources of stress most often cited by school-age children and adolescents." (OECD, 2015). So, the anticipation of academic failure can give students a threatening feel, resulting in test anxiety.

TA has many correlations on a child which are considered to be negative. Examples of factors correlating with TA are lack of parental engagement (Currie et al., 2012) and teacher engagement (Hoferichter, Raufelder & Eid, 2014), lower scores on intelligence tests (Von der Embse et al., 2018), more perceived bullying (Caputo, 2014), less sense of belonging in the classroom (Hoferichter et al., 2014; Steynmayr et al., 2018), and worse exam performance (Burnham & Makienko, 2018). Lower levels of motivation, increased general anxiety, negative self-beliefs, increased tension, a decline in attention span and a decline in concentration span are also factors associated with TA. Physical health complaints can also be correlated with TA (Currie et al., 2004). In extreme cases, TA has even a strong correlation with suicide (Pourtaleb, Mirnasab & Hadidi, 2018). Overall, girls have a significantly higher chance of getting TA than boys (Von der Embse et al., 2018).

Aims of the study

TA is a serious problem among students across the world. Even in the Netherlands, where the amount of TA under students is low compared to other countries (Currie et al, 2012; OECD, 2015), the percentage of 15-year-olds that feel pressure from school is over 40 per cent (Stevens et al., 2018). The correlators of TA are known, but there seems to be a knowledge gap in an international comparative context. This study aims to see which structures correlate with TA and if this differs per country. If the country does play a significant part, the external validity of existing literature could be less than thought.

As a case, two countries will be picked: one with a relatively low amount of TA (the Netherlands) and one with a relatively high amount of TA (Australia). This study will show what the relations are between the country on the one hand, and the correlators of TA and TA itself on the other hand. OECD's comparative study Programme for International Student Assessment [PISA] (OECD, 2015) will be used.

Empirical and Theoretical Background

In this part, relevant theories and empirical studies will be discussed regarding TA and its predictors. After this empirical and theoretical background, a conceptual model and a research question with logical hypotheses will be studied.

The OECD (2015) describes a model (see Figure 1) of measuring student well-being. They state that there are four closely related domains. Each dimension can be an outcome from the other sources or be an enabling condition for the other sources. The environment of the student is also essential. Environmental factors like school, family and teachers interact with the student and therefore significantly influence their well-being. Besides that, the contextual sources influence the proximal sources. For example, the school environment can be changed if governmental education policies are changed, which in turn changes the perceived students' well-being. In a healthy system, the three sources should be interdependent from each other. The students' well-being should depend on all other factors and not one factor (OECD, 2015).
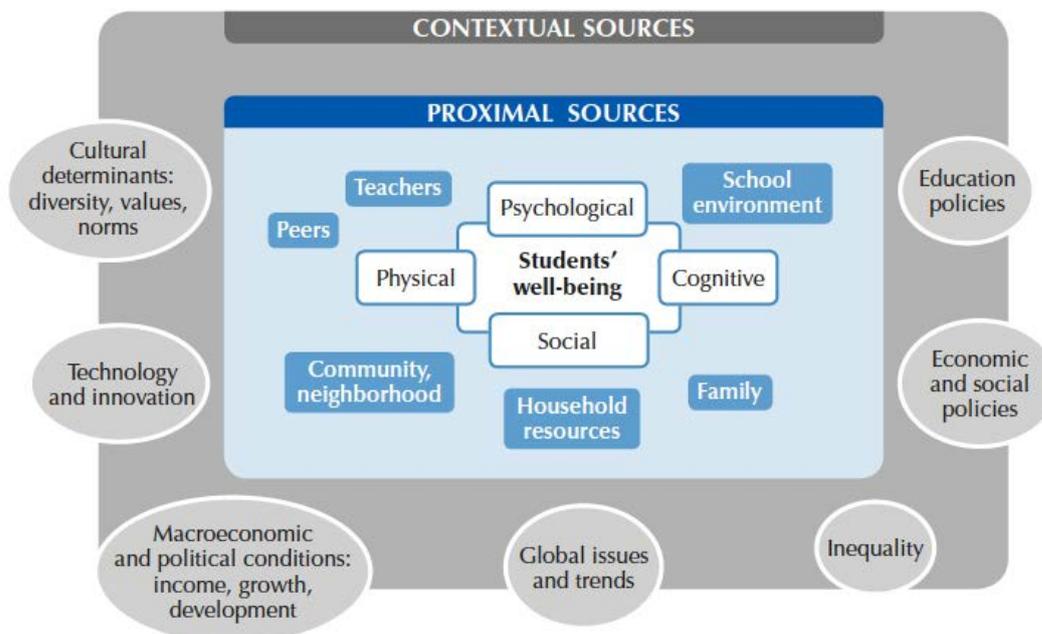

Figure 1. Dimensions and sources of students' well-being. Retrieved from PISA 2015 Results – Students' Well-Being (OECD, 2015).

Empirical research

Support from teachers has a positive influence on school-related behaviour of the child (Hoferichter et al., 2014). The connectedness that students with a high perceived teacher support feel will lead to less school-related stress. In their study, Hoferichter et al. (2014) asked

high school students about the perceived stress level, TA, and the level of perceived motivation from teachers and the level of the perceived teacher-student relationship. In their model, when teachers are positive motivators for the students, the students' TA levels are generally lower. Those students also maintain their interests in social and academic pursuits. According to Assor et al. (2005), girls generally are more accommodated to be more compliant against teachers and to be less assertive and more accommodative. Therefore, girls have more engagement with the teacher than boys.

Not only teachers are essential in the classroom. Hoferichter et al. (2014) describe the importance of inter-student relationships. In their piece of research, they found a negative relation between TA and support from peers. Wei and Chen (2010) found that a pupil's peer engagement and sense of belonging have a positive effect on their well-being. More sense of belonging would lead to lower delinquency and aggression, and higher motivation for achievement. A positive sense of belonging also has a positive effect on school enjoyment and school satisfaction and higher self-esteem. Steynmayr et al. (2018) state that the social warmth of the classroom is negatively correlated with both the emotionality and worry components of TA. According to Newman, Loman and Newman (2007), girls value peer support more than boys and feel more identified with their peers than boys.

A negative way of peer contact has also been researched. The study of Caputo (2014) states that TA and bullying are heavily correlated. Bullying would lead to negative attitudes toward school and anxiety about school failure and being evaluated. Lucas-Molina, Pérez-Albéniz and Fonseca-Pedrero (2018) studied the effect of bullying on well-being and suicidal thoughts and concluded that bullying indeed leads to lower well-being. The bullied peers take on the bad opinion about themselves from the bullies. The OECD (2015) also concludes that bullying and subjective well-being are negatively heavily correlated.

According to Brandmo, Bråten and Schewe (2018), ambition is also a positive correlator with TA. In their research, there was also a positive correlation between TA and being female. According to them, the social-cognitive theory of Bandura, personal goals are applauded, but a stronger commitment of a future career path predicts that there is more vulnerability in the evaluation of study results. When somebody with lots of ambition is failing a test, the cost is more than somebody with less ambition. According to Khalaila (2015), girls have more ambition, with as a result higher grades and more TA.

One of the predictors of TA in the literature is gender. Female students are more emotion-oriented and are more ambitious (Khalaila, 2015) and have more commitment to study (Caputo, 2014). It seems that the direct family expect more from girls than from boys (Brandmo,

Bråten & Schewe, 2018), and therefore have more TA than boys. Additionally, when girls do have TA, it is stronger for them than for boys (Brandmo, Bråten and Schewe, 2018). Von der Embse et al. (2018) concluded in their literature review that girls have a higher amount of TA than boys. This is the case in all years of education, but the difference in TA between the genders is lower in tertiary education than in primary- and secondary schools.

The model described in Figure 1 can be integrated well into the empirical research mentioned. TA (a psychological factor) is explained from the "proximal sources" in the model. Hoferichter et al. (2014) also state that students' well-being and TA have a positive relationship. This statement also corresponds with the model explained in Figure 1; the psychological factors of well-being (in this case TA) can interact with the well-being of students (OECD, 2015). As a social factor, the variables "teacher engagement" and "bullying" are mentioned. As a cognitive factor, the variable "ambition" is used. According to the OECD (2015), the cognitive factor says something about the foundation that students need to participate fully in society.

Conceptual model

A conceptual model can be made explaining which factors affect the TA of a student. However, some factors lack in this model. For example, well-being, which is a negative correlator with TA (Hoferichter et al., 2014) and predictors of TA (Brandmo, Bråten & Schewe, 2018; Steynmayr, 2018; Khalaila, 2015; Caputo, 2014), is not available in the data set for Australian students. Secondly, contextual factors, which play a significant role in the theoretical model described in Figure 1 (OECD, 2015), cannot be measured on an individual level. As a result, it would be guessing which contextual factors are affected in the role of TA and predictors of TA between various countries. Nuffic (2018a) describes the Australian education system as a system in which states influence the programme that high school students follow. A test decides admission to tertiary education at the end of high school. Only the best students in Australia can go to university. In the Netherlands, however, students choose earlier on which kind of tertiary education they want to follow. Finishing high school with pass grades is sufficient to go to the education of choice (Nuffic, 2018b).

One of the proximal factors that is not yet researched in empirical research is the physical factor. The OECD (2015) states that they do not measure physical factors in their study, except for self-reported information.

In Figure 2, the conceptual model of this study is shown. According to literature, TA is correlated with ambition (Brandmo, Bråten & Schewe, 2018), sense of belonging in the classroom (Steynmayr et al., 2018; Newman et al., 2007), perceived bullying (Caputo, 2014)

and teacher engagement (Hoferichter et al., 2014). However, the role of the country is unknown, due to a knowledge gap in the literature. The direct role of environmental factors of the country and the mediated role through the correlators will be examined through mediation analysis. It is expected that country is a direct correlator with TA, due to contextual factors as described in Figure 1.

The control variable in the model will be gender. It is expected that it correlates with TA (Khalaila, 2015; Von der Embse et al., 2018). It is expected, however, that the other control variables age and grade are not correlated with any of the variables. The subjects of this study are generally in the same grade and are fifteen years old.
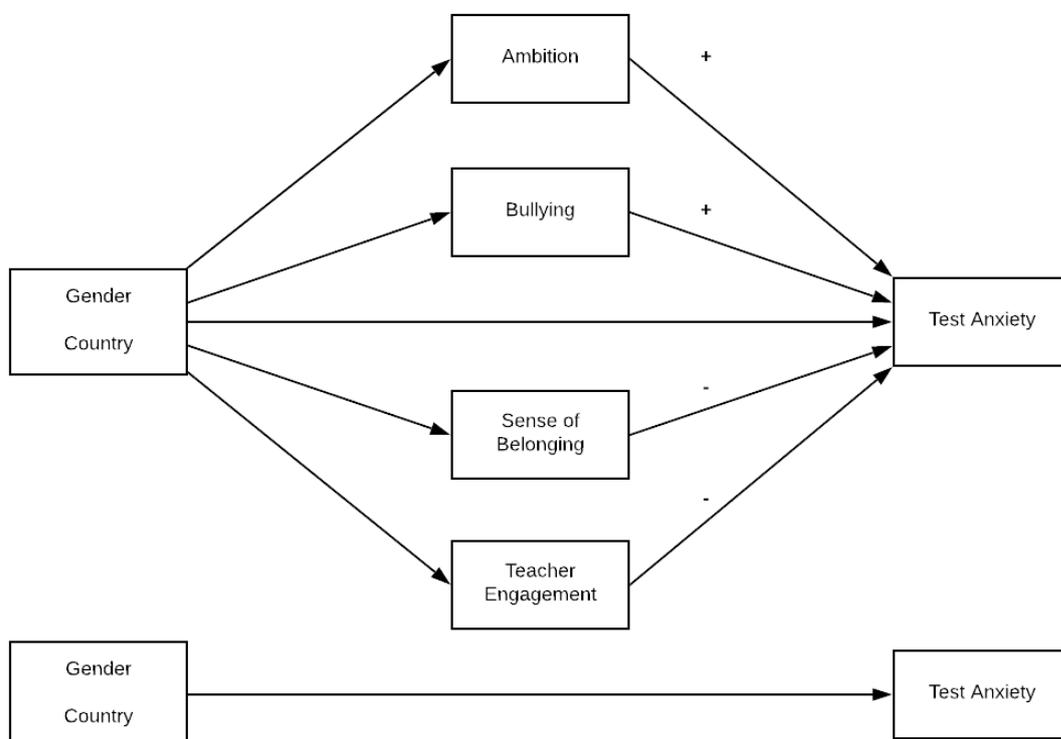

Figure 2. Conceptual model that will be used for a comparative study between Australia and the Netherlands.

This leads to the following research question and hypotheses.

Research question: How does test anxiety among 15-year old students differ between Australia and the Netherlands, and what is the role of their ambition, teachers and peers?

Q1: To what extent does a student's ambition mediate the association between country and levels of test anxiety?

According to the literature of Brandmo, Bråten and Schewe (2018), more ambition for higher grades will lead to more TA. It is expected that country is also a significant factor, as the contextual factors between Australia and the Netherlands differ (OECD, 2015; Figure 1).

- It was hypothesised that there was a partial mediation between country and TA, with ambition as mediator.

Q2: To what extent does perceived bullying mediate the association between country and levels of test anxiety?

The study of Caputo (2014) states that more perceived bullying will lead to more TA. It is expected that this as also the case in the current study. It is expected that country is also a significant factor, as the contextual factors between Australia and the Netherlands differ (OECD, 2015; Figure 1).

- It was hypothesised that there was a partial mediation between country and TA, with perceived bullying as mediator.

Q3: To what extent does perceived sense of belonging in the classroom mediate the association between country and levels of test anxiety?

Steynmayr et al. (2018) describe that more sense of belonging in the classroom will lead to less TA. It is expected that this as also the case in the current study. It is expected that country is also a significant factor, as the contextual factors between Australia and the Netherlands differ (OECD, 2015; Figure 1).

- It was hypothesised that there was a partial mediation between country and TA, with perceived sense of belonging in the classroom as mediator.

Q4: To what extent does teacher engagement mediate the association between country and levels of test anxiety?

According to the study of Hoferichter et al. (2014), higher engagement of teachers will lead to lower TA. Assor et al. (2005) also say that girls are socialised to be more accommodative to teachers, and therefore have more teacher engagement. It is expected that this as also the case

in the current study. It is expected that country is also a significant factor, as the contextual factors between Australia and the Netherlands differ (OECD, 2015; Figure 1).

- It was hypothesised that there was a partial mediation between country and TA, with teacher engagement as mediator.

Q5: To what extent is there a difference in levels of test anxiety between students in Australia and the Netherlands, mediated for ambition, bullying, sense of belonging in the classroom and teacher engagement?

From the literature, several correlators of test anxiety are mentioned. However, it is not known what the direct effect of country of origin is on TA, and what the effect of the correlators is on both TA and country.

- It was hypothesised that there was a direct effect between country and TA, with ambition, bullying, sense of belonging in the classroom and teacher engagement as mediators.

Method
Participants

This research was a secondary analysis from quantitative data that was gathered in 72 countries in 2015. More than 500,000 students participated in the Programme for International Student Assessment (PISA), a triennial worldwide research aimed to evaluate education systems worldwide (OECD, 2016). The OECD has a target population of students between 15 years and three months and 16 years and two months of age, and that are enrolled in an educational institution with grade 7 or higher. This study will only use data as collected among students in two countries: the Netherlands and Australia. See Table 1 for the target populations and samples for Australia and the Netherlands.

Table 1. *Target populations and samples for Australia and the Netherlands.*

Country	Total student population of 15-year olds	# of participating students	% student response rate
Australia	282 547	14 530	80.61%
The Netherlands	200 976	5 385	85.26%
Total	*484 523*	*19 915*	*81.84%*

Notes. Source: OECD (2016). Overall response rate states which percentage of the initial selected students finished all questionnaires. Total describes the total of Australia and The Netherlands.

Procedure

The sampling procedure for PISA was two-folded (OECD, 2016). Firstly, government representatives sampled individual schools in their countries where 15-year old students could be enrolled in. This happened systematically, intending to represent the most students. Per country, at least 150 schools were selected. If an institution declined, another institution was randomly selected. Secondly, within those schools, 42 students were randomly selected. Schools had to have an 80% participation rate for the students, measured nationally. When this percentage was too low, follow-up sessions were planned within schools with too few students participating. In the schools, the sampled students filled in the questionnaires and did the assignments they had to do in one day. For most countries, the Netherlands and Australia included, the questionnaires were to be completed on the computer. The time it took per student to fill in everything was approximately 3 hours and 30 minutes. The students could opt-out for the PISA survey by not showing up. Experts translated the questionnaires and filled in on the computer, by students, teachers and parents of the students. The current study focused on students, so only the students' questionnaire was used.

Instruments

The questions in the PISA study (OECD, 2015) about TA, ambition, teacher engagement in the classroom, sense of belonging in the classroom and bullying, along with gender and country, was used. See Table 6 in Appendix A for the variables and their codes in the PISA 2015

questionnaire. All items that form a scale in the PISA questionnaire were made a scale in this study, by using the mean scores of all items in a grade. The internal reliabilities were all in accordance with testing standards, differentiating between .829 and .927. The PISA questionnaire can be found at http://www.oecd.org/pisa/test/.

Independent variables (demographic variables)

The independent variables that are used in the study are country and gender. Both are nominal dichotomous variables which are represented by dummy variables. For gender, the reference category is female and for country, the reference category is Australian. These two variables are the only demographical variables in the study. Other demographical variables that are available in the data set, are not suited for the study. For example, it is improbable that the variables grade and age play a significant role in this study. After all, all students are approximately the same age and there is also not much variation in the students' grades.

Mediating variables

The mediating variables "ambition", "teacher engagement", "sense of belonging" and "bullying" will be used. All these variables exist of four or more questions, brought together in a scale. All mediating variables are rated on a four-point Likert scale from 1 (Strongly agree) to 4 (Strongly disagree). The scores of each participant gets centred, which means that the student has a negative number when they score under the mean, and vice-versa.

The scale "Exposure to bullying" exists of six questions about the perceived exposure to bullying that the student experiences. The student can choose between "Never or almost never", "A few times a year", "A few times a month" or "Once a week or more". An example of a question in this scale is "Other students made fun of me".

The scale "Ambition" exists of five questions about the student's ambition. An example of a question is "I want to be one of the best students in my class".

The scale "Teacher engagement in the classroom" exists of five questions about the perceived support students receive from their science teacher. An example of a question is "The teacher helps students with their learning".

The scale "Sense of belonging in the classroom" exists of six questions about how students feel regarding other students in the classroom. An example of a question is "I feel like I belong at school".

Outcome variable

The scale "Test anxiety" exists of five questions about the perceived TA rate of the students, rated on a Likert scale from 1 (Strongly disagree) to 4 (Strongly agree). Each participant gets a mean score ranging from 1 to 4 from those five items. An example of a question is "I get very

tense when I study for a test". The Cronbach's alpha reliabilities for this scale are .852 for Australia and .833 for the Netherlands.

The programme used to analyse the data is R (R Core Team, 2013). The analyses that were used for the research questions were mediation, multiple regression and ANOVA. Both the multiple regression analysis and the mediation analysis were tested on normality. All conditions were met to use a mediation analysis for the research questions. See Appendix B for the R syntax and results of the normality tests.

Chi-square- and t-tests were used to look for differences between the two countries within a variable. After that, a multiple regression model was used to see what the correlations between country, gender and TA were. Then, a mediation model for gender and country as independent variables, and TA as the dependent variable, and each of the mediators apart has been run. In the end model, all mediators were in the model, together with the independent variables country and gender, and the dependent variable TA.

Modules used in R were tidyverse (Wickham, 2017), haven (Wickham & Miller, 2019), psych (Revelle, 2018), Hmisc (Harrell Jr, 2019), DAAG (Maindonald & Braun, 2015), Broom (Robinson & Hayes, 2018) and e1071 (Meyer et al, 2019).

Results

For optimal results, only participants that filled in all questions about TA are included. This keeps 18 819 of the initial 19 915 students (94,50% of all students). Of those 18 819 students, 13 695 (72,8%) are Australian and 5 124 (27,2%) are Dutch. Demographics of these participants are shown in Table 2. In Appendix B, the results of the assumptions are shown.

Table 2. *Descriptive analyses of variables used in the study.*

	Australia		*Netherlands*		*Total*		*t-test/chi^2-test between*
	Mean/%	*SD*	*Mean/%*	*SD*	*Mean/%*	*SD*	*countries*
Demographics							
Gender: Female	49,8%		50,7%		50,0%		$\chi^2 = 1.36$
Country: Australian					72,2%		
Scale scores							
Ambition#	0,13	0,56	-0,35	0,48	0	0,58	t = 52,72***
Bullying#	0,07	0,58	-0,18	0,32	0	0,54	t = 37,51***
Sense of belonging#	0,05	0,58	-0,14	0,52	0	0,57	t = 23,14***
Teacher engagement#	0,04	0,77	-0,10	0,69	0	0,76	t = 10,84***
Dependent variable							
Test Anxiety	2,70	0,67	2,15	0,60	2,55	0,70	t = 53,98***

Note. #: These variables are centred, with the mean value of all participants at 0. The p-value is measured before centring the X variables.
*: p<.05, **: p<.01, ***: p<.001.

In Table 3, a regression matrix is shown of the independent variables used in the model. As shown, variables generally are correlated with each other, except for bullying and teacher engagement. However, no correlation is r > 0,8.

Table 3. *Correlations between various independent variables.*

	Ambition	Bullying	Sense of belonging	Teacher engagement
Ambition	-	.04***	-.08***	.12***
Bullying	.04***	-	.38***	.01
Sense of Belonging	-.08***	.38***	-	-.07***
Teacher engagement	.12***	.01	-.07***	-

Note. *** p<.001. Numbers in table corresponds to the Pearson's r value.

The relation between country and gender as independent variables and TA as dependent variable, without mediating variables, is shown in Table 4. A positive correlation is found between TA and the demographical factors (R^2 =.18, F(2,18816)=2071.9, p<.001).

Table 4. *Multiple regression model with dependent variable TA and demographical factors.*

	Beta estimate	t-value	p-value
Intercept	1.981	197.74	<.001
Female	0.331	35.23	<.001
Australia	0.558	35.83	<.001

Note. *** p<.001. Numbers in table corresponds to the Pearson's r value.

When ambition is added as a mediator, a partial mediation exists in the relationship between country and TA, mediated for ambition (total effect β=0.56, t=53.78, p<.001; direct effect β=0.51, t=46.26, p<.001; indirect effect β=0.05, CI [0.04,0.06]). For the relationship between gender and ambition, there is a positive correlation for girls (β=0.33, t=35.83, p<.001) but no mediation effect. The total model is highly significant (R^2 =.19, F(3,18815)=1450.27, p<.001). There is also a positive correlation between ambition and TA (β=0.11, t=13.07, p<.001). See Figure 3 for the model.
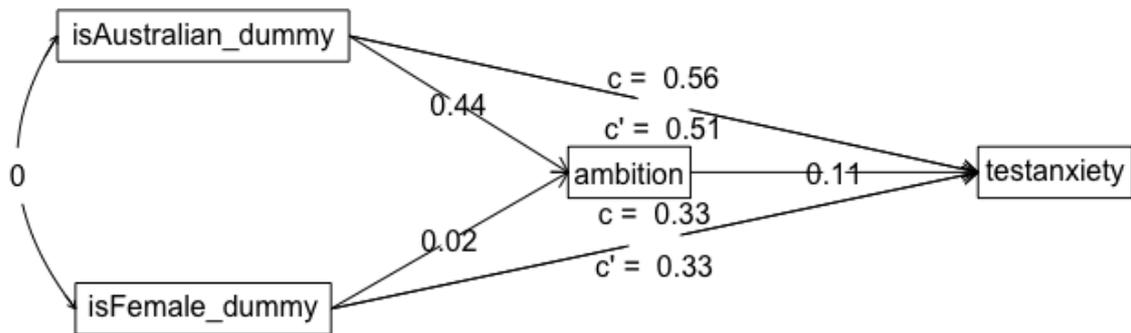
Figure 3. Mediation model with ambition as mediating variable.

When bullying is a mediator, there is a partial mediation between country and TA, mediated for bullying (total effect β=0.56, t=53.78, p<.001; direct effect β=0.50, t=48.20, p<.001; indirect effect β=0.05, CI [0.05,0.05]). For the relationship between gender and TA, mediated by bullying, there is a positive correlation between being a girl and TA, but not mediated via bullying (direct effect β=0.33, t=35.83, p<.001). The total model is highly significant (R^2 =.21, F(3,18815)=1632.32, p<.001). There is also a positive correlation between bullying and TA (β=0.21, t=24.87, p<.001). See Figure 4 for the model.
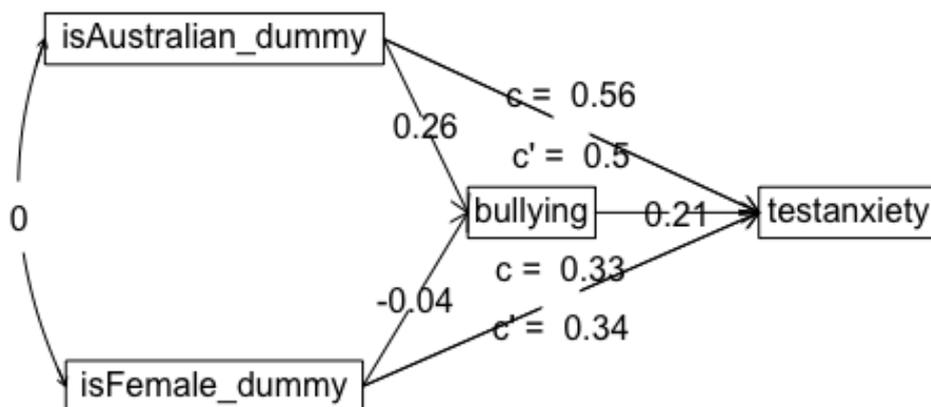

Figure 4. Mediation model with bullying as mediating variable.

When sense of belonging in the classroom is a mediator between country and TA, the correlation between country and TA is partially mediated by sense of belonging (total effect β=0.56, t=53.78, p<.001 direct effect β=0.51, t=49.26, p<.001 indirect effect β=0.05, CI [0.04,

0.05]). For the relationship between gender and TA, there is a positive correlation between being a girl and TA, but not via the mediating factor sense of belonging (direct effect β=0.31, t=34.36, p<.001). The whole model is highly significant (R^2 =.21, F(3,18815)=1676.43, p<.001). There is also a positive correlation between sense of belonging and TA (β=0.2, t=26.96, p<.001). See Figure 5 for the model.
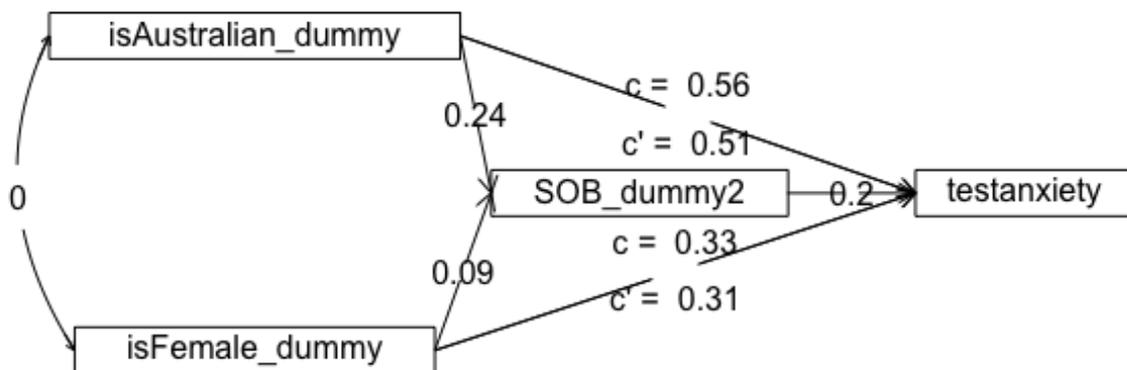
Figure 5. Mediation model with sense of belonging as mediating variable.

Teacher engagement plays no mediating factor between country and TA (total effect β=0.56, t=53.78, p<.001 direct effect β=0.56, t=53.84, p<.001 indirect effect β=0, CI [0,0]). However, the effect between sense of belonging and TA is significant (t=-2.87, p=.004). Sense of belonging is also no mediator between gender and TA (total effect β=0.33, t=35.83, p<.001 direct effect β=0.33, t=35.15, p<.001 indirect effect β=0, CI [0,0.01]). The whole model is significant (R^2 =.18, F(3,18815)=1384.12, p<.001). See Figure 6 for the model.
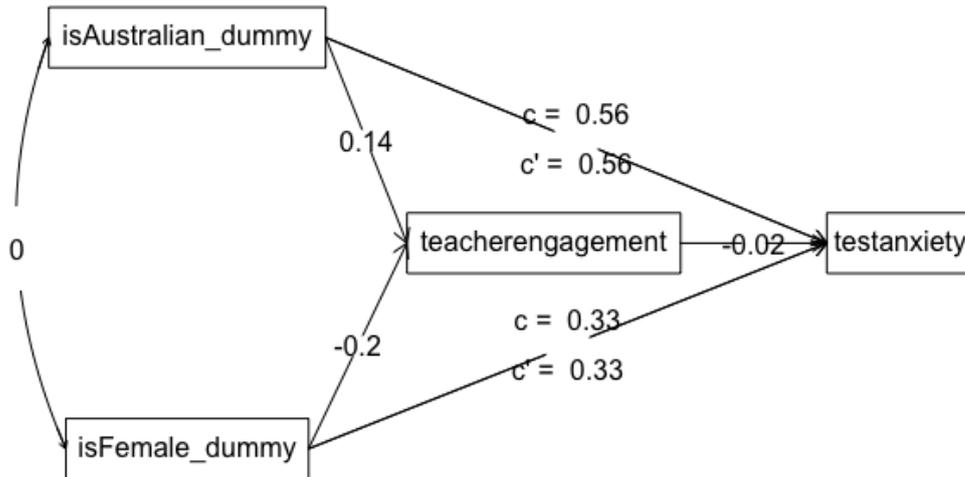
Figure 6. Mediation model with teacher engagement as mediating variable.

A model can be made which tries to explain if the country is correlated with TA, with three of the four predictors of TA as mediating factors, and gender as a control variable. Figure 7 displays this model, and in Table 5, the statistics of this model are shown. For country, there is a partial mediation, through ambition, bullying, and sense of belonging. For gender, there is no mediation, only a direct positive correlation between being female and TA. The whole model is highly significant ($R^2 = .24$, $F(5,18813)=1201.47$, $p<.001$).

Figure 7. Full mediation model.

Table 5. *Mediation effects for the correlation between gender, country and TA, mediated for three variables.*

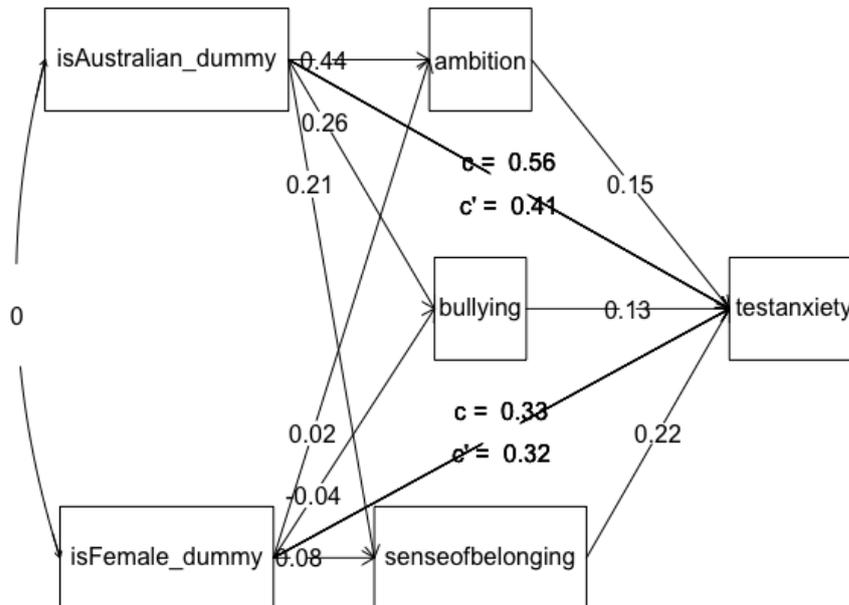| Effect from country	b	t	Lower	Upper
			95% CI	
Total	0.56	53.78***		
Direct	0.41	37.50***		
Indirect (mediation)	0.15		0.14	0.16
Effect from gender	b	t	Lower	Upper
Total	0.33	35.83***		
Direct	0.32	35.44***		
Indirect (mediation)	0.01		0.01	0.02
Indirect effects for each mediator	b	t	Lower	Upper
Country:				
Ambition	0.07		0.06	0.08
Bullying	0.03		0.03	0.04
Sense of belonging	0.05		0.04	0.05
Gender:				
Ambition	0.00		0.00	0.00
Bullying	-0.01		-0.01	0.00
Sense of belonging	0.02		0.01	0.02

Note. * = p<.05; ** = p<.01; *** = p<.001.

Discussion

This paper tried to find predictors to explain test anxiety under Dutch and Australian children. Summarising this paper, it is shown that country and gender are highly correlated in the prevalence of TA under 15-year old children. A part of this correlation can be explained by three of the four mediating factors: ambition, sense of belonging in the classroom and perceived bullying. However, the final mediation model only has four per cent more explained variance than the model with only demographical factors. This means that gender and country are still the largest correlators with TA.

Ambition (RQ1) plays a significant mediating role in the correlation between TA and country, where being Australian is positively correlated with TA. However, ambition is not mediated for the correlation between gender and TA. There is also a significant positive relationship in the model between TA and ambition. However, the direct effect of country on TA is larger than the effect of ambition alone on TA. The fact that ambition and TA are correlated is expected from the literature (Brandmo, Bråten & Schewe, 2018). It is also expected from the literature that environmental factors could cause differences in country regarding TA. The study shows that country is indeed a significant correlator with TA. The insignificance

between gender and ambition, however, is not aligned with the literature; according to Khalaila (2015), girls should have more ambition than boys. They feel more pressure to succeed academically from family (Brandmo, Bråten & Schewe, 2018) and thus feel more pressure to get higher grades than boys. The literature takes ambition into account as a mediating factor to explain the positive correlation between TA and being a girl, but this model does not. One explanation of this difference is that the contextual factors of the literature are different from this study. In this study, Dutch and Australian students are compared. The sample of Brandmo, Bråten and Schewe (2018) are Norwegian and Khalaila (2015) compares Jewish and Arabic children. It could be that Dutch and Australian girls experience less pressure from their family to perform well.

Perceived bullying (RQ2) plays a significant mediating role in the correlation between TA and country. Being Australian is positively correlated with TA. However, bullying itself is not mediated for the correlation between gender and TA. There is a correlation between gender and TA, however. The fact that perceived bullying and TA are positively correlated is indeed explained in the literature. Caputo (2014) states that bullying leads to a negative self-evaluation in the classroom. Victims of bullying feel worthless, and that leads to negative motivation and attitudes for school, and therefore, those victims have higher general anxiety and TA. Caputo (2014) states that boys are bullied more than girls, and therefore, girls would have less TA than boys. However, in the analysis, there is no direct correlation between gender and bullying. This also can be explained from environmental factors: Caputo's study (2014) exists from only Italian students. The role that environmental factors play can be different between Italian, Dutch and Australian students.

Perceived sense of belonging in the classroom (RQ3) plays a significant mediating role in the correlation between TA and country. Being Australian is positively correlated with TA. This effect is for country mediated by sense of belonging in the classroom: Australian children have more sense of belonging in the classroom and therefore have more TA. However, this is not what the literature states. Steynmayr et al. (2018) state that sense of belonging is negatively correlated with TA. However, in this study, sense of belonging is positively correlated with TA. Newman et al. (2007) also say that females have more sense of belonging in the classroom than males. This is not the case in this study. According to the model, the fact that females suffer from TA is not mediated through sense of belonging in the classroom. One explanation, again, could be that the literature is not Australian or Dutch-based, but from the United States.

Perceived teacher engagement (RQ4) does not play a significant role in the relationship between country and TA. Teacher engagement is not a mediating factor, and also the direct correlation between teacher engagement and TA is barely significant. This is strange because, from the literature, it is expected that teacher engagement and TA are negatively correlated (Hoferichter et al., 2014). This is the case, but the correlation is barely significant. Gender does correlate with teacher engagement. This is what literature also states (Assor et al., 2005). Being Australian is also correlated with teacher engagement, which can be explained by environmental factors (OECD, 2015). It is difficult to find a reason why perceived teacher engagement behaves so differently in this model than in the literature. Contextual factors could play a role: being Australian or Dutch is different from being Israeli or German, where the literature concentrates.

In the full model (RQ5), both gender and country do have a direct and indirect effect on TA. This can be explained by combining the previous research questions (except teacher engagement). Most of the expectations from the literature are met: only the effect between gender and bullying (not correlated, although Caputo (2014) states that boys are bullied more), and gender and ambition (not correlated, although Khalaila (2015) states that girls have more ambition than boys) are not met. However, the full model states clearly that even with gender as a control variable, there is a substantial significant direct effect between country and TA. This full model explains 24 per cent of the variance, which is a medium-large effect. A model with only gender, country and TA explains 18 per cent of the variance, which is a medium effect.

The mediation model (see Figure 6; Table 5) explains 24 per cent of the variance regarding TA. The role of gender, implemented as a control variable, is quite significant. Being a female causes TA to be 0.32 points higher (on a scale from 1 to 4). Mediating factors explain 0.01 point in the correlation between gender and TA. It is quite strange that the mediating factors have almost no effect on gender and TA. The past literature has proven otherwise (Assor et al., 2005; Khalaila, 2015; Newman et al., 2007).

The most significant impact of the independent variables on TA is country. Australian students score 0.41 point higher than Dutch students on TA (on a scale from 1 to 4), and with the mediation model, it is even 0.56 point higher. It is still the question why this variance between Dutch and Australian students is so high. One explanation could be the role of environmental factors. The Nuffic (2018a; 2018b) states that the school structure of Australia and the Netherlands is quite different: for example, in the Netherlands, a division is made between pre-university, pre-vocational, and senior general secondary education when the students are 11 or

12. When the students finish their education (this is around 17 or 18 years), they can apply to their respective tertiary pathway (Nuffic, 2018a). However, in Australia, this division is not made. The Australian students all do the same test, and only a top percentage can apply for university. This test can decide a student's further career, and it can be a significant factor in the stress level of an Australian student.

Another possible explanation could be that the students took the PISA survey in May, where it was spring in the Netherlands but autumn in Australia (OECD, 2016). The weather could play a part here, where people feel better and less anxious when it is sunnier and warmer (Howarth & Hoffman, 1984). A third explanation could be the absence of various other predictors of TA. One predictor that is not weighed in, but is expected to correlate with TA (Hoferichter et al., 2014), is well-being. Unfortunately, data about students' well-being are not available for Australian students. Additional correlators with TA that can be taken into account (Ringeisen & Raufelder, 2015), are parental support and parental pressure.

A strength of this study is that this study tries to explain TA through various mediating factors, and sees what effect the factors, and country itself, has on TA. This is also the first study that tries to find a model explaining TA by taking contextual factors into account. This study took the conceptual model used by the OECD (2015; see Figure 1) and tried to replicate all relevant factors, including contextual factors. However, that is more difficult than it seems. In the PISA questionnaire (2015), no questions are asked about contextual factors. This is a limitation of the study. Another limitation is the fact that this study is cross-sectional. This implies that only correlation can be shown. It is also not possible to deduce direction in the correlation. The third limitation is the fact that only 15-year old students participate in the study. Australian students have important tests on that age (Nuffic, 2018a), and it is possible that the results would be different when this study looked at students at another age.

In the PISA questionnaire (OECD, 2015), Australia scored high on perceived TA and low on sense of belonging, while the Netherlands scored vice-versa. 71.9% of Australian children say that they have a feeling that they belong at school, against 80.9% of Dutch children. Moreover, compared to the Netherlands (29.7%), more Australian students (74.2%) report that they are ambitious and want to be one of the best students in their class (OECD, 2015). The OECD (2015) also states that bullying in Australia (24.2%) is more frequent than in The Netherlands (9.3%). However, in the literature, no explanation is given why this difference exists. The model does also not explain why there is a massive difference in TA, and the predictors of TA, between the Netherlands and Australia. Future research should study the differences between various countries (and contextual factors, like educational systems and

national policy on education) to see what causes this variation in TA, and what the implications are for existing literature in not taking country into account. For example, Spain, Italy and Japan each have a higher TA than Australia (OECD, 2015). Future research can look at those countries to see what explains the variation in TA between those countries and other countries whose students have lower TA. This research can also be further expanded by taking into account Australian states and various school types in Australia and the Netherlands. In Appendix B, some preliminary results are shown.

Concluding, it is still unknown where the gap in TA between Australian and Dutch students lies. Gender and some of the correlators of TA do play a role, but further research has to be made to see which contextual factors, and which other factors, lie underneath this huge difference between the two countries.

References

Assor, A., Kaplan, H., Kanat-Maymon, Y, & Roth, G. (2005). Directly controlling teacher behaviors aspredictors of poor motivation and engagementin girls and boys: The role of anger and anxiety. *Leaning and Instruction, 15,* 397-413. doi:10.1016/j.learninstruc.2005.07.008

Brandmo, C., Bråten, I., & Schewe, I. (2017). Social and personal predictors of test anxiety among Norwegian secondary and postsecondary students. *Social Psychology of Education, 22(43), 43-61*. doi:10.1007/s11218-018-9461-y

Burnham, T. A., & Makienko, I. (2018). Factors Affecting Exam Completion Speed, Exam Performance, and Nonexam Performance. *Journal of Marketing Education, 40(2),* 140-151. doi:10.1177/0273475317715290

Caputo, A. (2014). Psychological Correlates of School Bullying Victimization: Academic Self-Concept, Learning Motivation and Test Anxiety. *International Journal of Educational Psychology, 3(1),* 69-99. doi:10.4471/ijep.2014.04

Currie, C., Roberts, C., Morgan, A., Smith, R., Settertobulte, W., Samdal, O., & Rasmussen, V.B. (2004). *Young people's health in context. Health Behaviour in School-aged Children (HBSC) study: international report from the 2001/2002 survey*. Retrieved from http://www.euro.who.int/__data/assets/pdf_file/0008/110231/e82923.pdf.

Currie, C., Zanotti, C., Morgan, A., Currie, D., de Looze, M., Roberts, C., Samdal, O., Smith, O. R.F., & Barnekow, V. (2012). *Social determinants of health and well-being among young people.* Copenhagen: WHO Regional Office for Europe.

Harrell Jr, F. E. (2019). *Hmisc: Harrell Miscellaneous.* Retrieved from https://CRAN.R-project.org/package=Hmisc

Howarth, E., & Hoffman, M. S. (1984). A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology, 75(1)*, 15-23. doi:10.1111/j.2044-8295.1984.tb02785.x

Hoferichter, F., Raufelder, D., & Eid, M. (2014). The mediating role of socio-motivational relationships in the interplay of perceived stress, neuroticism, and test anxiety among adolescent students. *Psychology in the Schools, 51(47),* 736-752. doi:10.1002/pits.21778

Khalaila, R. (2015). The relationship between academic self-concept, intrinsic motivation, test anxiety, and academic achievement among nursing students: Mediating and

moderating effects. *Nurse Education Today, 35,* 432-438. doi: 10.1016/j.nedt.2014.11.001

Lawrence, D., Johnson, S., Hafekorst, J., Boterhoven de Haan, K., Sawyer, M., Ainley, J., & Zubrick, J. R. (2015). *The Mental Health of Children and Adolescents. Report on the second Australian Child and Adolescent Survey of Mental Health and Wellbeing.* Canberra, Australia: Department of Health.

Lucas-Molina, B., Pérez-Albéniz, A., & Fonseca-Pedrero, E. (2018). The potential role of subjective wellbeing and gender in the relationship between bullying or cyberbullying and suicidal ideation. *Psychiatry Research, 270,* 595-601. doi:10.1016/j.psychres.2018.10.043

Maindonald, J.H., & Braun, J. (2015). DAAG: Data Analysis and Graphics Data and Functions. Retrieved from https://CRAN.R-project.org/package=DAAG.

Meyer, D., Dimitriadou, E., Hornik, E., Weingessel, A., & Leisch, D. (2019). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071).* Austria: Vienna. Retrieved from https://CRAN.R-project.org/package=e1071.

Newman, B.M., Lohman, B.J., & Newman, P.R. (2007). Peer group membership and a sense of belonging: their relationship to adolescent behavior problems. *Adolescence, 42(166),* 1-24.

Nuffic. (2018a). *Education system Australia described and compared with the Dutch system.* The Hague, the Netherlands. Retrieved from https://www.nuffic.nl/en/subjects/education-and-diplomas-australia/

Nuffic. (2018b). *Education system The Netherlands.* The Hague, the Netherlands. Retrieved from https://www.nuffic.nl/en/subjects/education-and-diplomas-the-netherlands/

Organisation for Economic Cooperation and Development (OECD). (2015). Programme for International Student Assessment (Volume III). OECD Publishing, Paris. doi:10.1787/9789264273856-en

Organisation for Economic Cooperation and Development (OECD). (2016). PISA 2015 Technical Report. OECD Publishing, Paris.

Pourtaleb, N., Mirnasab, M., & Hadidi, Y. (2018). The Effectiveness of Integrated Training Programme (ITP) in Decreasing Female Students' Test Anxiety: Moderating Effect of Mothers' Perfectionism. *Educational Psychology in Practice, 34(4)*, 386-396. doi:10.1080/02667363.2018.1468734

Putwait, D., Chamberlain, S., Daly, A.L., & Sadreddini, S. (2014). Reducing test anxiety among school-aged adolescents: a field experiments. *Educational Psychology in Practice, 30(4),* 240-220. doi:10.1080/02667363.2014.964392

R Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research.* Northwestern University, USA: Evanston, Illinois, USA. Retrieved from https://CRAN.R-project.org/package=psych

Ringeisen, T., & Raufelder, D. (2015). The interplay of parental support, parental pressure and test anxiety - Gender differences in adolescents. *Journal of Adolescence, 45,* 67-75. doi:10.1016/j.adolescence.2015.08.018

Robinson, D., & Hayes, A. (2018). broom: Convert Statistical Analysis Objects into Tidy Tibbles. Retrieved from https://CRAN.R-project.org/package-broom.

Stevens, G., van Dorsselaer, S., Boer, M., de Roos, S., Duinhof, E., ter Bogt, T., van den Eijnden, R., Kuyper, L., Visser, D., Vollebergh, W., & de Looze, M. (2018). *HBSC 2017: Gezondheid en Welzijn in Nederland.* Utrecht, the Netherlands: Utrecht University.

Steynmayr, R., Heyder, A., Neumburg, C., Michels, J., & Wirthwein, L. (2018). School-Related and Individual Predictors of Subjective Well-Being and Academic Achievement. *Frontiers in Psychology, 9(2631),* 1-16. doi:10.3389/fpsyg.2018.02631

UNICEF. (2013). *Child Well-being in rich countries. A comparative overview.* Retrieved from UNICEF website: https://www.unicef-irc.org/publications/pdf/rc11_eng.pdf.

Von der Embse, N., Jester, D., Roy, D., & Post. J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483-493. doi:10.1016/j.jad.2017.11.048

Wei, H., & Chen, J. (2010). School attachment among Taiwanese adolescents: the roles of individual characteristics, peer relationships, and teacher well-being. *Social Indicators Research, 95(3),* 421-436. doi:10.1007/s11205-009-9529-3

Wickham, H. (2017). *tidyverse: Easily Install and Load the 'Tidyverse'.* Retrieved from https://CRAN.R-project.org/package=tidyverse.

Wickham, H., & Miller, E. (2019). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.* Retrieved from https://CRAN.R-project.org/package=haven.

World Health Organization. (2017). *Depression and Other Common Mental Disorders. Global Health Estimates.* Geneva, Switzerland: World Health Organization.

Appendix A: Overview of variables used in the study.

Table 6. *Overview of variables and question codes used in the study. Note: Source: OECD (2015); the questions with a * will be scored reversely.*

Question code	Question
CNT	Country code 3-character
ST004D01T	Student (Standardized) Gender.
STRATUM	Stratum ID 7-character.[1]
ST034Q01TA*	I feel like an outsider (or left out of things) at school.
ST034Q02TA	I make friends easily at school.
ST034Q03TA	I feel like I belong at school.
ST034Q04TA*	I feel awkward and out of place in my school.
ST034Q05TA	Other students seem to like me.
ST034Q06TA*	I feel lonely at school.
ST038Q03NA	Other students left me out of things on purpose.
ST038Q04NA	Other students made fun of me.

---

[1] In the Netherlands, the stratum ID distinguishes secondary education levels (vmbo/havo/vwo). In Australia, the stratum ID distinguishes governmental, independent and catholic schools per state.

ST038Q05NA	I was threatened by other students.
ST038Q06NA	Other students took away or destroyed things that belonged to me.
ST038Q07NA	I got hit or pushed around by other students.
ST038Q08NA	Other students spread nasty rumours about me.
ST104Q01NA	The teacher tells me how I am performing in this course.
ST104Q02NA	The teacher gives me feedback on my strengths in this subject.
ST104Q03NA	The teacher tells me in which areas I can still improve.
ST104Q04NA	The teacher tells me how I can improve my performance.
ST104Q05NA	The teacher advises me on how to reach my learning goals.
ST118Q01NA	I often worry that it will be difficult for me taking a test.
ST118Q02NA	I worry that I will get poor grades at school.
ST118Q03NA	Even if I am well prepared for a test, I feel very anxious.
ST118Q04NA	I get very tense when I study.
ST118Q05NA	I get nervous when I don't know how to solve a task at school.
ST119Q01NA	I want top grades in most or all of my courses.

ST119Q02NA	I want to be able to select from among the best opportunities available when I graduate.
ST119Q03NA	I want to be the best, whatever I do.
ST119Q04NA	I see myself as an ambitious person.
ST119Q05NA	I want to be one of the best students in my class.

Appendix B: R Script for running statistical analyses

Assumptions

Here are the assumptions of the mediation analysis.

```
#Model: Y ~ Xs + gender + country

multreg_GOOD <- lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement +CNT + GENDER, data = dataset_All)

multregNoCNT <- lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement +GENDER, data = dataset_All)

multregOnlyCountry <- lm(testanxiety ~ CNT + GENDER, data = dataset_All)

#Model summary of good multiple regression
summary(multreg_GOOD)
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER, data = dataset_All)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.22964 -0.38219  0.01997  0.40250  2.25492
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.977892   0.012394 159.580   <2e-16 ***
## ambition           0.135485   0.009828  13.785   <2e-16 ***
## bullying           0.169197   0.010454  16.186   <2e-16 ***
## SOB_dummyMedium    0.210473   0.011310  18.609   <2e-16 ***
## SOB_dummyHigh      0.263605   0.022984  11.469   <2e-16 ***
## teacherengagement -0.014757   0.007002  -2.107   0.0351 *
## CNTAustralia       0.406443   0.012928  31.438   <2e-16 ***
## GENDERFemale       0.311811   0.010504  29.685   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6097 on 13835 degrees of freedom
##   (4976 observations deleted due to missingness)
```

```
## Multiple R-squared:  0.2357, Adjusted R-squared:  0.2353
## F-statistic: 609.5 on 7 and 13835 DF,  p-value: < 2.2e-16
#Assumption 1: linearity per X variable

#Plot for Ambition per country
ggplot(data = NULL, show.legend = TRUE) +
 geom_smooth(mapping = aes(x = dataset_AUS$ambition, y = dataset_AUS$testanxiety), color ="#00008B") +
 geom_smooth(mapping = aes(x = dataset_NL$ambition, y = dataset_NL$testanxiety), color ="Orange") +
 geom_smooth(mapping = aes(x = dataset_All$ambition, y = dataset_All$testanxiety), color ="Black") +
 labs(fill = "Country", name = "Ambition and TA per country", labels=c("Australia","The Netherlands"), x = "Ambition scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
#Skewness of Ambition
skewness(dataset_All$ambition, na.rm = TRUE)
## [1] -0.2177902
#Bullying graph
ggplot(data = NULL, show.legend = TRUE) +
```

```
geom_smooth(mapping = aes(x = dataset_AUS$bullying, y = dataset_AUS$testanxiety), color ="#00008B") +
geom_smooth(mapping = aes(x = dataset_NL$bullying, y = dataset_NL$testanxiety), color ="Orange") +
geom_smooth(mapping = aes(x = dataset_All$bullying, y = dataset_All$testanxiety), color ="Black") +
labs(name = "Bullying and TA per country", labels=c("Australia","The Netherlands"), x ="Bullying scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
#Skewness of Bullying
skewness(dataset_All$bullying, na.rm = TRUE)
## [1] 2.429485
#Old Sense of Belonging in the classroom plot
ggplot(data = NULL, show.legend = TRUE) +
```

```
geom_smooth(mapping = aes(x = dataset_AUS$senseofbelonging, y = dataset_AUS$testanxiety), color = "#00008B") +
geom_smooth(mapping = aes(x = dataset_NL$senseofbelonging, y = dataset_NL$testanxiety), color = "Orange") +
geom_smooth(mapping = aes(x = dataset_All$senseofbelonging, y = dataset_All$testanxiety), color = "Black") +
labs(name = "Sense of Belonging and TA per country", labels=c("Australia","The Netherlands"), x = "Sense of Belonging scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```
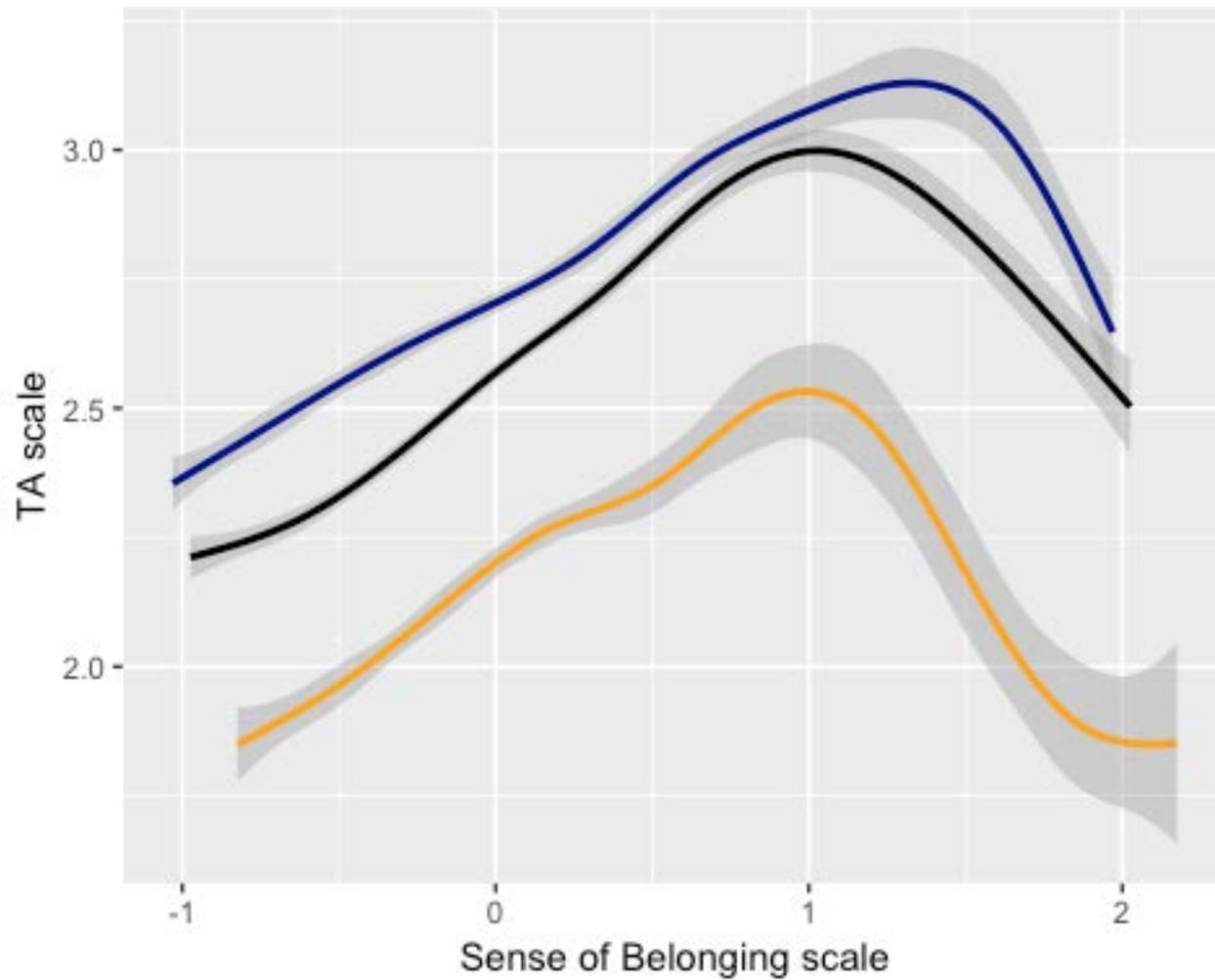
```
#Skewness of Sense of Belonging
skewness(dataset_All$senseofbelonging, na.rm = TRUE)
## [1] 0.8554281
#SOB_Dummy interaction plot
#Interaction plot SOB x CNT
```

```r
interaction.plot(x.factor = dataset_All$SOB_dummy, trace.factor = dataset_All$CNT, response =dataset_All$testanxiety, xlab = "Sense of
Belonging", ylab = "Test Anxiety", trace.label ="Country")
```



```r
#Teacher Engagement regression plot
ggplot(data = NULL, show.legend = TRUE) +
 geom_smooth(mapping = aes(x = dataset_AUS$teacherengagement, y =dataset_AUS$testanxiety), color = "#00008B") +
```

```r
 geom_smooth(mapping = aes(x = dataset_NL$teacherengagement, y = dataset_NL$testanxiety), color = "Orange") +
 geom_smooth(mapping = aes(x = dataset_All$teacherengagement, y = dataset_All$testanxiety), color = "Black") +
 labs(name = "Teacher engagement and TA per country", labels=c("Australia","The Netherlands"), x = "Teacher engagement scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```
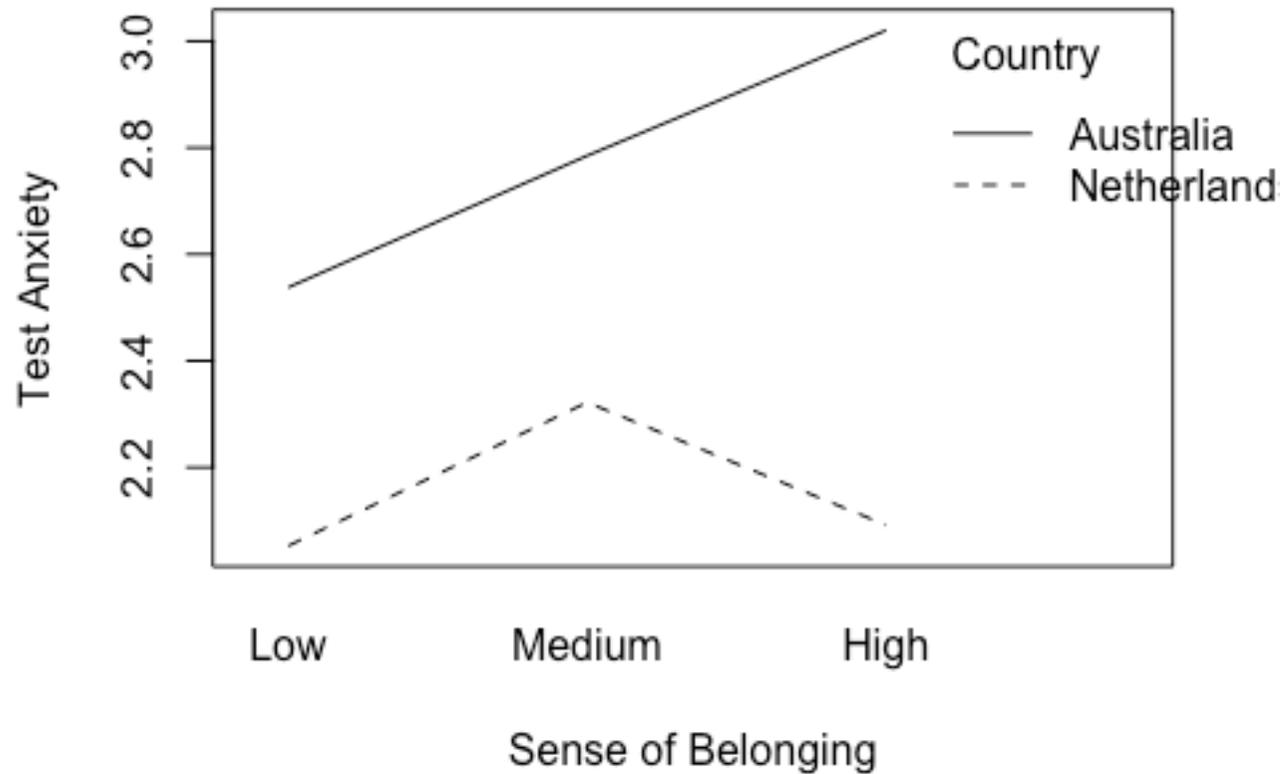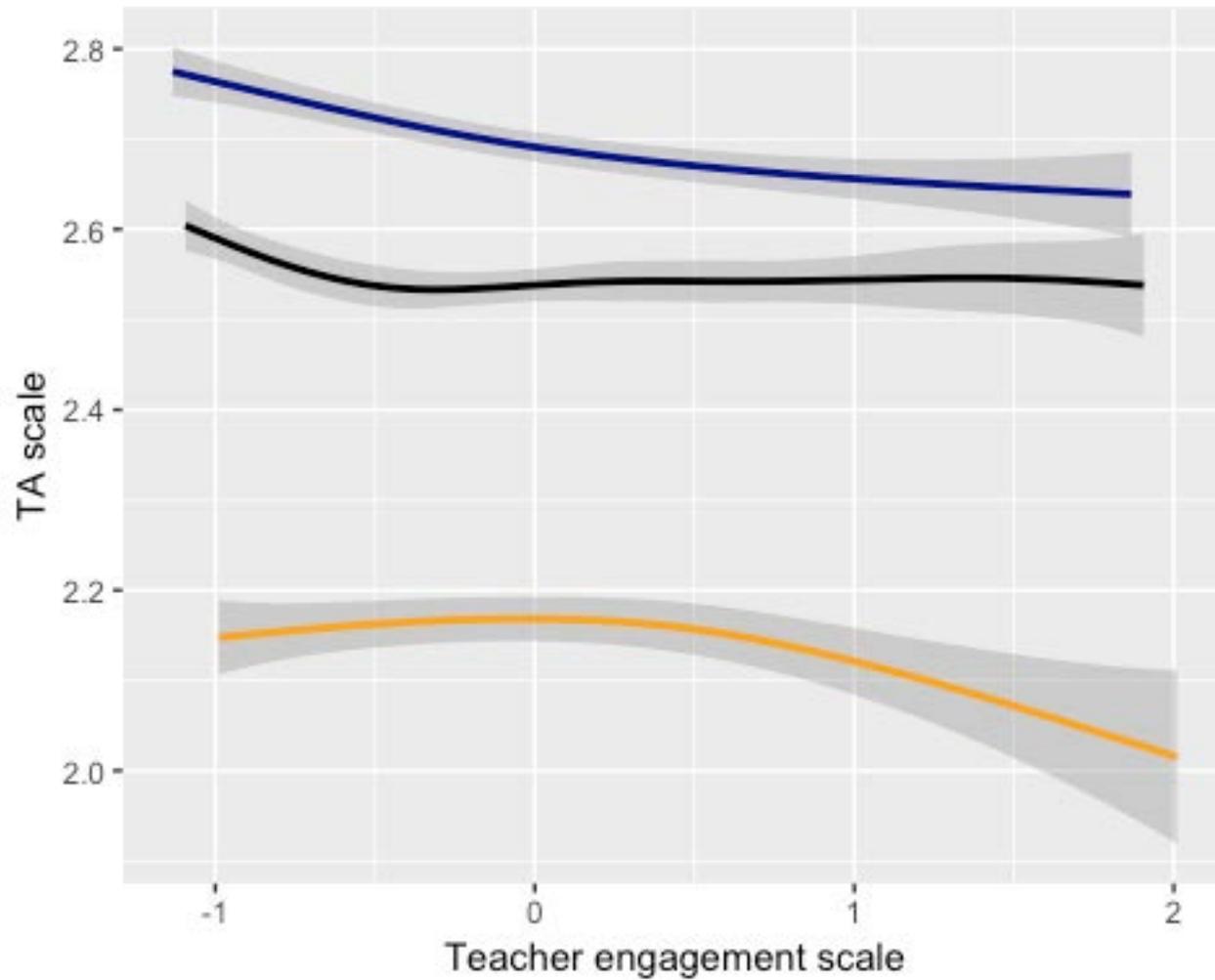
```
skewness(dataset_All$teacherengagement, na.rm = TRUE)
## [1] 0.4303407
#Assumption 2: Correlation Table between Key Xs

dplyr::select(dataset_All, ambition, bullying, senseofbelonging, teacherengagement) %>%
```

```
 as.matrix() %>%
 rcorr(type = c("pearson"))
##                   ambition bullying senseofbelonging teacherengagement
## ambition              1.00     0.04            -0.08              0.12
## bullying              0.04     1.00             0.38              0.01
## senseofbelonging     -0.08     0.38             1.00             -0.07
## teacherengagement     0.12     0.01            -0.07              1.00
##
## n
##                   ambition bullying senseofbelonging teacherengagement
## ambition             18280    17743            17724             14390
## bullying             17743    18212            17741             14449
## senseofbelonging     17724    17741            18174             14357
## teacherengagement    14390    14449            14357             14709
##
## P
##                   ambition bullying senseofbelonging teacherengagement
## ambition                    0.0000   0.0000           0.0000
## bullying          0.0000             0.0000           0.0932
## senseofbelonging  0.0000   0.0000                     0.0000
## teacherengagement 0.0000   0.0932   0.0000
```
#Variance Influence Factor of multiple regression
```
vif(multreg_GOOD)
##         ambition          bullying    SOB_dummyMedium      SOB_dummyHigh
##           1.1994            1.1656             1.1873             1.1811
## teacherengagement       CNTAustralia       GENDERFemale
##           1.0412             1.2491             1.0272
```
#Standard residuals of cases
```
augment(multreg_GOOD) %>%
 ggplot() +
 geom_histogram(mapping = aes(x = .std.resid)) +
 labs(name = "Standard residuals", y = "Number of cases", x = "Standard residuals of model")
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
#QQplot
plot(multreg_GOOD, 2)
```

Normal Q-Q

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

```
#Residuals VS Fitted plot
plot(multreg_GOOD, 1)
```

Residuals vs Fitted

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

```
# Number of standard residuals >3 in multiple regression
augment(multreg_GOOD) %>%
 filter(.std.resid > 3 | .std.resid < -3) %>%
 dplyr::select(.std.resid)
```

42
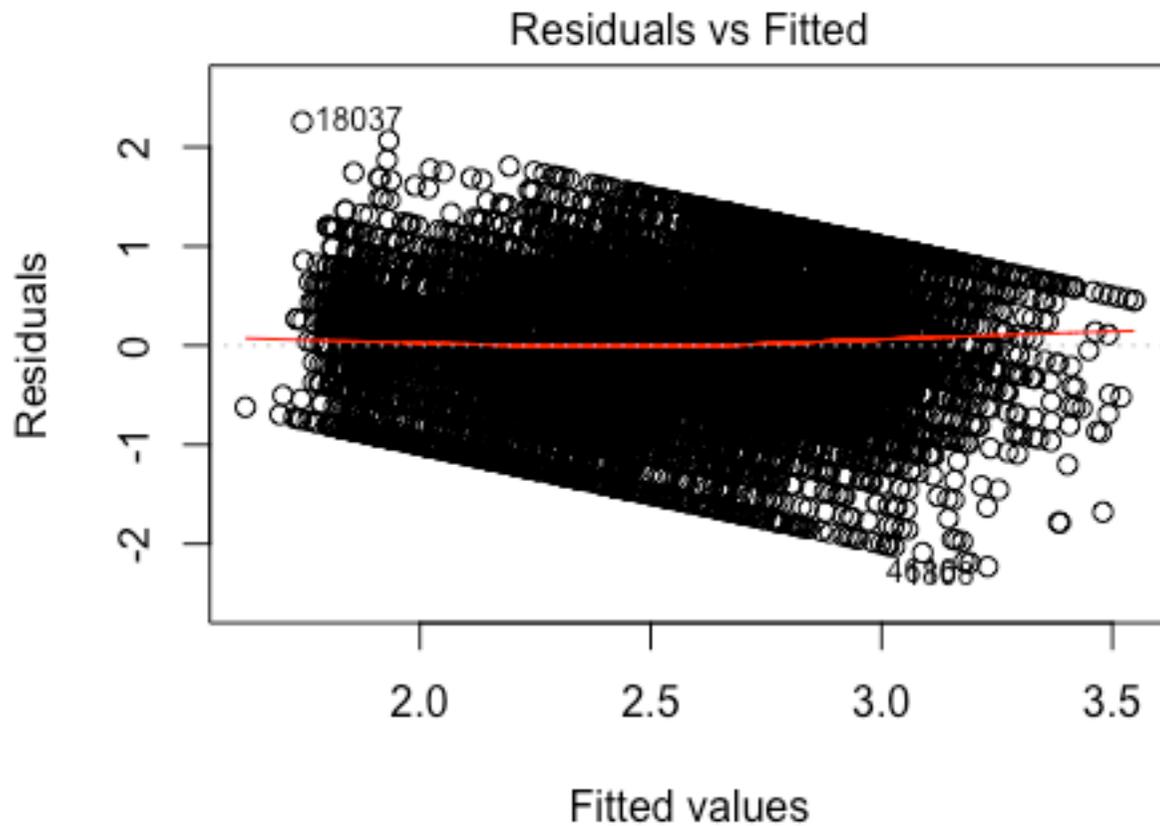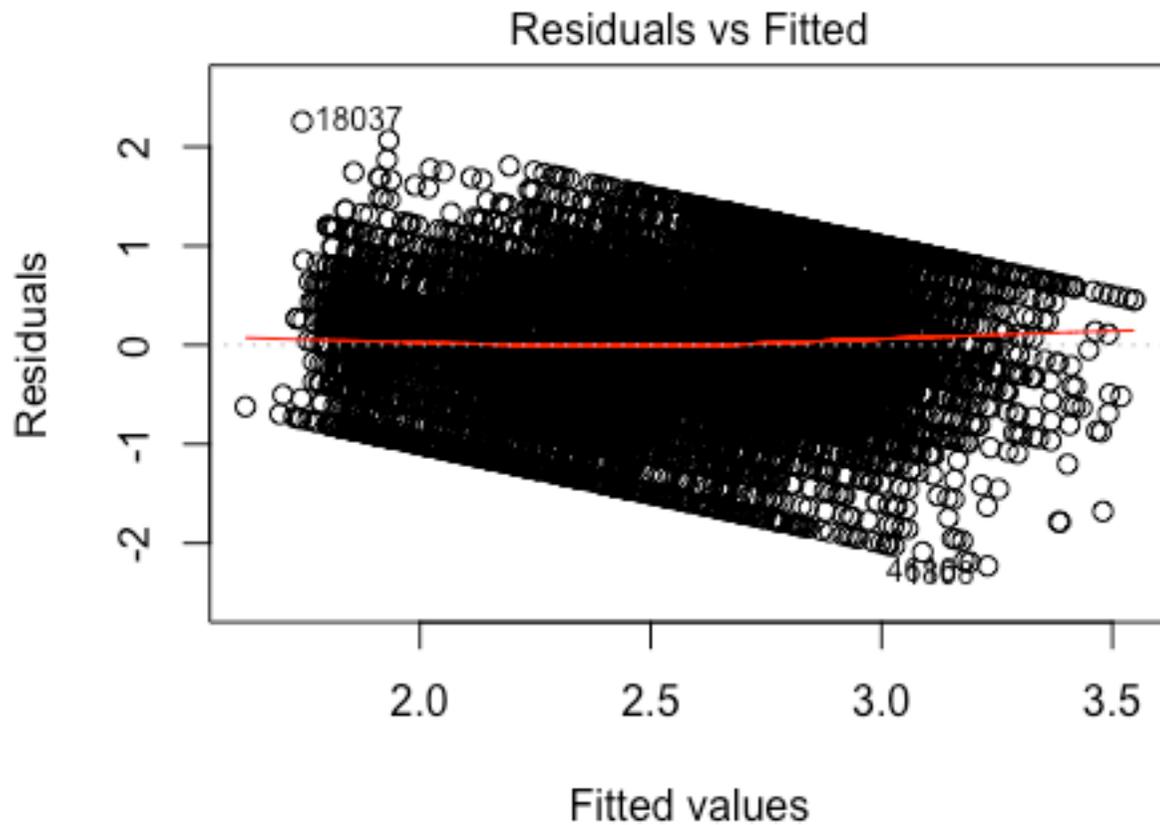
```
## # A tibble: 28 x 1
##    .std.resid
##         <dbl>
##  1      -3.43
##  2      -3.02
##  3      -3.22
##  4      -3.30
##  5      -3.01
##  6      -3.17
##  7      -3.66
##  8      -3.25
##  9      -3.20
## 10      -3.02
## # … with 18 more rows
#Cooks Distance
summary(augment(multreg_GOOD)$.cooksd)
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.000e+00 5.250e-06 2.488e-05 8.221e-05 8.483e-05 4.895e-03
#Leverage plot
plot(multreg_GOOD, 5)
```

Residuals vs Leverage

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

State and school types

In this chunk, linear regression models of state and school types are worked out.

#Question - Does TA in various Australian states or school types differ?

```
#Multiple regression in Australia with differences in states
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER + State, data = dataset_AUS) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##      teacherengagement + GENDER + State, data = dataset_AUS)
##
## Residuals:
##      Min      1Q   Median       3Q      Max
## -2.17151 -0.39031  0.01765  0.41638  1.73067
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.473173   0.028173  87.785  < 2e-16 ***
## ambition                 0.129141   0.011434  11.295  < 2e-16 ***
## bullying                 0.160532   0.011488  13.974  < 2e-16 ***
## SOB_dummyMedium          0.208861   0.014232  14.675  < 2e-16 ***
## SOB_dummyHigh            0.231107   0.030861   7.489 7.55e-14 ***
## teacherengagement       -0.019624   0.008194  -2.395   0.0166 *
## GENDERFemale             0.345001   0.012653  27.266  < 2e-16 ***
## StateNew South Wales    -0.016449   0.029781  -0.552   0.5807
## StateNorthern Territory -0.024369   0.043857  -0.556   0.5785
## StateQueensland         -0.002420   0.031178  -0.078   0.9381
## StateSouth Australia    -0.032250   0.032258  -1.000   0.3174
## StateTasmania           -0.082321   0.035677  -2.307   0.0211 *
## StateVictoria           -0.061771   0.031142  -1.984   0.0473 *
## StateWestern Australia  -0.040355   0.031945  -1.263   0.2065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6214 on 9982 degrees of freedom
##   (3699 observations deleted due to missingness)
## Multiple R-squared:  0.1392, Adjusted R-squared:  0.1381
## F-statistic: 124.2 on 13 and 9982 DF,  p-value: < 2.2e-16
#Multiple regression in Australia with school types
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER +SchoolType, data = dataset_AUS) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
```

```
##     teacherengagement + GENDER + SchoolType, data = dataset_AUS)
##
## Residuals:
##      Min      1Q   Median      3Q     Max
## -2.18561 -0.38566  0.01889  0.41558  1.73496
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.470050   0.014800 166.890  < 2e-16 ***
## ambition           0.129206   0.011419  11.315  < 2e-16 ***
## bullying           0.160981   0.011491  14.009  < 2e-16 ***
## SOB_dummyMedium    0.210865   0.014247  14.800  < 2e-16 ***
## SOB_dummyHigh      0.231577   0.030863   7.503 6.75e-14 ***
## teacherengagement -0.019887   0.008176  -2.432  0.01502 *
## GENDERFemale       0.344053   0.012652  27.193  < 2e-16 ***
## SchoolTypePrivate -0.054518   0.018695  -2.916  0.00355 **
## SchoolTypePublic  -0.030512   0.015297  -1.995  0.04612 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6214 on 9987 degrees of freedom
##   (3699 observations deleted due to missingness)
## Multiple R-squared:  0.1386, Adjusted R-squared:  0.1379
## F-statistic: 200.9 on 8 and 9987 DF,  p-value: < 2.2e-16
#Multiple regression in the Netherlands with school type
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER +STRATUM, data = dataset_NL) %>%
  summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER + STRATUM, data = dataset_NL)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7334 -0.3568  0.0129  0.3446  2.1906
##
## Coefficients:
##                                Estimate Std. Error t value
## (Intercept)                     1.88185    0.01921  97.979
## ambition                        0.11656    0.01964   5.934
## bullying                        0.20239    0.03032   6.675
```

46

```
## SOB_dummyMedium                          0.22521    0.01957  11.510
## SOB_dummyHigh                             0.10138    0.04400   2.304
## teacherengagement                        0.00792    0.01378   0.575
## GENDERFemale                             0.23532    0.01876  12.547
## STRATUMNetherlands - HAVO or VWO         0.07248    0.01900   3.815
## STRATUMNetherlands - Private Education  0.17775    0.57633   0.308
##                                         Pr(>|t|)
## (Intercept)                             < 2e-16 ***
## ambition                                3.22e-09 ***
## bullying                                2.83e-11 ***
## SOB_dummyMedium                         < 2e-16 ***
## SOB_dummyHigh                           0.021261 *
## teacherengagement                       0.565636
## GENDERFemale                            < 2e-16 ***
## STRATUMNetherlands - HAVO or VWO        0.000139 ***
## STRATUMNetherlands - Private Education 0.757775
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5758 on 3838 degrees of freedom
##   (1277 observations deleted due to missingness)
## Multiple R-squared:  0.09506,    Adjusted R-squared:  0.09318
## F-statistic:  50.4 on 8 and 3838 DF,  p-value: < 2.2e-16
```

Appendix C: Full R Script

Master Thesis about test anxiety under 15-year old Dutch and Australian students

This explorative study tries to understand the predictors of Test Anxiety under 15-year old Dutch and Australian students. With four predictors (Ambition, Bullying, Sense of Belonging in the Classroom and Teacher Engagement) and four confounders (Grade, Gender, Age and Country), I try to predict Test Anxiety under those students. The data set used is the Programme for International Student Assessment (PISA) 2015 cohort, with only the Dutch and Australian students. Goal of this study is to fill in the country-comparative knowledge gap within Test Anxiety and the predictors of Test Anxiety.

Setup

This chunk of code is the setup for the analysis. Packages will be installed and the 2015 PISA dataset will be downloaded from the website. Only relevant variables and participants are included.

```r
#Install packages
install.packages(c("tidyverse","haven","psych","Hmisc","DAAG","broom"), repos ="http://cran.us.r-project.org") #Download packages.

#Open packageswar
library(tidyverse) #This package includes ggplot2 (to make graphs) and dplyr (data manipulation).
library(haven) #This package can read from and write to SPSS-files.

#Open two data sets in R

temp <- tempfile()
download.file(url = "https://webfs.oecd.org/pisa/PUF_SPSS_COMBINED_CMB_STU_QQQ.zip", destfile = temp) #Download PISA data in temporary folder

dataset_all_main_variables <- unz(temp, "CY6_MS_CMB_STU_QQQ.sav") %>% #Extract dataset 1 from temporary folder
 read_sav() %>% #Import dataset 1 in R
 tbl_df() %>% #Make Tibble (better version of a data frame)
 dplyr::select(CNTSTUID,CNT,ST004D01T,ST034Q01TA,ST034Q02TA,ST034Q03TA,ST034Q04TA,ST034Q05TA,ST034Q06TA,ST104Q01NA,ST104Q02NA,ST104Q03NA,ST104Q04NA,ST104Q05NA,ST118Q01NA,ST118Q02NA,ST118Q03NA,ST118Q04NA,ST118Q05NA,ST119Q01NA,ST119Q02NA,ST119Q03NA,ST119Q04NA,ST119Q05NA,STRATUM) %>% #Select only relevant variables
 rename(GENDER = ST004D01T) #Rename the "Gender"-variable to GENDER

dataset_all_extra_variables <- unz(temp, "CY6_MS_CMB_STU_QQ2.sav") %>% #Extract dataset 2 from temporary folder
 read_sav() %>% #Import dataset 2 in R
 tbl_df() %>% #Make Tibble (better version of a data frame)
 dplyr::select(CNTSTUID,ST038Q03NA,ST038Q04NA,ST038Q05NA,ST038Q06NA,ST038Q07NA,ST038Q08NA) #Select only relevant variables
```

```r
dataset_work <- merge(dataset_all_main_variables, dataset_all_extra_variables, by="CNTSTUID") %>% #Merge data sets
 filter(CNT == "NLD" | CNT == "AUS")  %>% #Select only Dutch and Australian students
 select(-CNTSTUID) #Remove student ID from the data set

remove(list = c("temp","dataset_all_main_variables","dataset_all_extra_variables")) #Remove unused datasets

write_sav(dataset_work, "sav-files/dataset_work.sav", compress = TRUE) #Make SPSS-file of relevant variables and cases.
```

Open packages

In this chunk, the packages needed for the analysis are opened.

```r
#Open packages
library(tidyverse) #This package includes ggplot2 (to make graphs) and dplyr (data manipulation).
library(haven) #This package can read from and write to SPSS-files.
library(psych) #Package for reverse coding and mediation
library(Hmisc) #For changing labels
library(DAAG) #For VIF measurements
library(broom) #Broom can be used to measure leverages and standardised dfs
library(e1071) #Package for measuring skewness
```

Data cleaning

In this chunk, the data will be cleaned. The SPSS-file that is saved in the previous chunk is loaded and reverse coded.

```r
#####OPEN DATASET#####
dataset_work <- read_sav("sav-files/dataset_work.sav") %>% #Import dataset in R.
 as_tibble() #Import as tibble, a better version of a data frame in R.

#ASSIGN LEVELS TO VARIABLES -- Factorisation is the way of making ordinal variables in R - the different numbers become factors with their respective
labels, but they are treated like ordinal variables.

dataset_work$CNT <- factor(dataset_work$CNT) #Make the Country variable nominal.

levels(dataset_work$CNT) <- c("Australia", "Netherlands") #Data in "Country" have to be converted to Netherlands/Australia (NLD = the Netherlands,
AUS = Australia)
dataset_work$STRATUM <- factor(dataset_work$STRATUM) #Data in 'STRATUM' has to be converted to their respective strata, as mentioned in the PISA 2015
Code Book.
levels(dataset_work$STRATUM) <- c("ACT-Catholic, Year 10","ACT-Catholic, No Year 10","ACT-Gov, Year 10","ACT-Gov, No Year 10","ACT-Ind, Year
10","NSW-Catholic, Year 10","NSW-Catholic, No Year 10","NSW-Gov, Year 10","NSW-Gov, No Year 10","NSW-Ind, Year 10","NSW-Ind, No Year 10","VIC-
Catholic, Year 10","VIC-Catholic, No Year 10","VIC-Gov, Year 10","VIC-Gov, No Year 10","VIC-Ind, Year 10","VIC-Ind, No Year 10","QLD-Catholic, Year
10","QLD-Catholic, No Year 10","QLD-Gov, Year 10","QLD-Gov, No Year 10","QLD-Ind, Year 10","QLD-Ind, No Year 10","SA-Catholic, Year 10","SA-Catholic,
NO Year 10","SA-Gov, Year 10","SA-Gov, No Year 10","SA-Ind, Year 10","SA-Ind, No Year 10","WA-Catholic, Year 10","WA-Gov, Year 10","WA-Ind, Year
```

```
10","TAS-Catholic, Year 10","TAS-Gov, Year 10","TAS-Gov, No Year 10","TAS-Ind, Year 10","NT-Catholic, Year 10","NT-Gov, Year 10","NT-Gov, No Year
10","NT-Ind, Year 10","Netherlands - PRO or VMBO","Netherlands - HAVO or VWO","Netherlands - Private Education") #Data in "Strata" have to be
converted to their respective labels, as seen in the PISA 2015 codebook.

dataset_work$GENDER <-  factor(dataset_work$GENDER) #Make Nominal Factors - All nominal variables will be converted to "Factors".
levels(dataset_work$GENDER) <- c("Female", "Male") #Data in "Gender" have to be converted to Male/Female (1=Female, 2=Male)

dataset_work <- dataset_work %>%
 mutate(CNT = relevel(CNT, ref = "Netherlands")) %>% #Reference of country is Dutch
 mutate(GENDER = relevel(GENDER, ref = "Male")) #Reference of gender is Male

#Transform variables to numeric
dataset_work <- transform(dataset_work, ST118Q01NA = as.numeric(ST118Q01NA),
                          ST118Q02NA = as.numeric(ST118Q02NA),
                          ST118Q03NA = as.numeric(ST118Q03NA),
                          ST118Q04NA = as.numeric(ST118Q04NA),
                          ST118Q05NA = as.numeric(ST118Q05NA),
                          ST034Q01TA = as.numeric(ST034Q01TA)
)

#Reverse code three variables in new data set
dataset_work2 <- reverse.code(c(-1,-1,-1), dplyr::select(dataset_work, ST034Q01TA,ST034Q04TA,ST034Q06TA),1,5) %>%
 as_tibble() %>%
 rename("ST034Q01TA"="ST034Q01TA-") %>%
 rename("ST034Q04TA"="ST034Q04TA-") %>%
 rename("ST034Q06TA"="ST034Q06TA-")
label(dataset_work2$ST034Q01TA) <- "(R) I feel like an outsider (or left out of things) at school."
label(dataset_work2$ST034Q04TA) <- "(R) I feel awkward and out of place in my school."
label(dataset_work2$ST034Q06TA) <- "(R) I feel lonely at school."

#Replace the old variables with the reversed variables
dataset_work <- dplyr::select(dataset_work, -c(ST034Q01TA,ST034Q04TA,ST034Q06TA)) %>%
 cbind(dataset_work2) %>%
 as_tibble()
remove("dataset_work2")
```

Subset questions

Subsets of the the predictors and Test Anxiety are made.

```
#Select subset of questions
TA <- dplyr::select(dataset_work, ST118Q01NA,ST118Q02NA,ST118Q03NA,ST118Q04NA,ST118Q05NA) %>%
```

```r
 na.omit()
Ambition <-dplyr::select(dataset_work,ST119Q01NA,ST119Q02NA,ST119Q03NA,ST119Q04NA,ST119Q05NA) %>%
 na.omit()
Belonging <-dplyr::select(dataset_work,ST034Q01TA,ST034Q02TA,ST034Q03TA,ST034Q04TA,ST034Q05TA,ST034Q06TA) %>%
 na.omit()
Bullying <-dplyr::select(dataset_work,ST038Q03NA,ST038Q04NA,ST038Q05NA,ST038Q06NA,ST038Q07NA,ST038Q08NA) %>%
 na.omit()
Teacher <-dplyr::select(dataset_work,ST104Q01NA,ST104Q02NA,ST104Q03NA,ST104Q04NA,ST104Q05NA) %>%
 na.omit()

#Make various new variables for all predictors and TA; those are means.

dataset_All <- mutate(dataset_work, senseofbelonging = (ST034Q01TA + ST034Q02TA +ST034Q03TA + ST034Q04TA + ST034Q05TA + ST034Q06TA)/6) %>%
 mutate(teacherengagement = (ST104Q01NA + ST104Q02NA + ST104Q03NA + ST104Q04NA +ST104Q05NA)/5) %>%
 mutate(testanxiety = (ST118Q01NA + ST118Q02NA + ST118Q03NA + ST118Q04NA +ST118Q05NA)/5) %>%
 mutate(ambition = (ST119Q01NA + ST119Q02NA + ST119Q03NA + ST119Q04NA +ST119Q05NA)/5) %>%
 mutate(bullying = (ST038Q03NA + ST038Q04NA + ST038Q05NA + ST038Q06NA +ST038Q07NA + ST038Q08NA)/6)

remove(dataset_work) #Remove dataset_work

#Only keep cases that do not have any TA missings
dataset_All <- dataset_All[complete.cases(dataset_All$testanxiety),]
```

Make country-specific subsets and questions

There are a few country-specific questions; for the Netherlands it is the school type (havo/vwo; and vmbo), and for Australia it is state and school type (public/private/catholic schools). Two subsets are made for each of the countries and those questions.

```r
#Make State level + school type level (Australian)

dataset_AUS_ACT
<- filter(dataset_All, as.numeric(STRATUM) == 1 | as.numeric(STRATUM) == 2 | as.numeric(STRATUM) == 3 | as.numeric(STRATUM) == 4 | as.numeric(STRATUM
) ==5) %>%
 add_column(State = "ACT")
dataset_AUS_NSW
<- filter(dataset_All, as.numeric(STRATUM) == 6 | as.numeric(STRATUM) == 7 | as.numeric(STRATUM) == 8 | as.numeric(STRATUM) == 9 | as.numeric(STRATUM
) ==10 | as.numeric(STRATUM) == 11) %>%
 add_column(State = "New South Wales")
dataset_AUS_VIC
<- filter(dataset_All, as.numeric(STRATUM) == 12 | as.numeric(STRATUM) == 13 | as.numeric(STRATUM) == 14 | as.numeric(STRATUM) == 15 | as.numeric(STR
ATUM) == 16 | as.numeric(STRATUM) == 17) %>%
 add_column(State = "Victoria")
```

```r
dataset_AUS_QLD
<- filter(dataset_All, as.numeric(STRATUM) == 18 |as.numeric(STRATUM) == 19 | as.numeric(STRATUM) == 20 | as.numeric(STRATUM) == 21 |as.numeric(STRATUM) == 22 | as.numeric(STRATUM) == 23) %>%
 add_column(State = "Queensland")
dataset_AUS_SA
<- filter(dataset_All, as.numeric(STRATUM) == 24 | as.numeric(STRATUM) == 25 | as.numeric(STRATUM) == 26 | as.numeric(STRATUM) == 27 | as.numeric(STRATUM) == 28 | as.numeric(STRATUM) == 29) %>%
 add_column(State = "South Australia")
dataset_AUS_WA <- filter(dataset_All, as.numeric(STRATUM) == 30 | as.numeric(STRATUM) == 31 | as.numeric(STRATUM) == 32) %>%
 add_column(State = "Western Australia")
dataset_AUS_TAS
<- filter(dataset_All, as.numeric(STRATUM) == 33 | as.numeric(STRATUM) == 34 | as.numeric(STRATUM) == 35 | as.numeric(STRATUM) == 36) %>%
 add_column(State = "Tasmania")
dataset_AUS_NT
<- filter(dataset_All, as.numeric(STRATUM) == 37 | as.numeric(STRATUM) == 38 | as.numeric(STRATUM) == 39 | as.numeric(STRATUM) == 40) %>%
 add_column(State = "Northern Territory")

dataset_AUS <- bind_rows(dataset_AUS_ACT, dataset_AUS_NSW, dataset_AUS_NT, dataset_AUS_QLD, dataset_AUS_SA, dataset_AUS_TAS, dataset_AUS_VIC, dataset_AUS_WA)

dataset_AUS_Gov
<- filter(dataset_AUS, as.numeric(STRATUM) == 3 | as.numeric(STRATUM) == 4 | as.numeric(STRATUM) == 8 | as.numeric(STRATUM) == 9 | as.numeric(STRATUM) ==14 | as.numeric(STRATUM) == 15 | as.numeric(STRATUM) == 20 | as.numeric(STRATUM) ==21| as.numeric(STRATUM) == 26 | as.numeric(STRATUM) == 27 | as.numeric(STRATUM) ==31| as.numeric(STRATUM) == 34 | as.numeric(STRATUM) == 35 | as.numeric(STRATUM) ==38| as.numeric(STRATUM) == 39) %>%
 add_column(SchoolType = "Public")
dataset_AUS_Private
<- filter(dataset_AUS, as.numeric(STRATUM) == 5 |as.numeric(STRATUM) == 10 | as.numeric(STRATUM) == 11 | as.numeric(STRATUM) == 16 |as.numeric(STRATUM) == 17 | as.numeric(STRATUM) == 22 | as.numeric(STRATUM) == 23 |as.numeric(STRATUM) == 28 | as.numeric(STRATUM) == 29 | as.numeric(STRATUM) == 32 | as.numeric(STRATUM) == 36 | as.numeric(STRATUM) == 40) %>%
 add_column(SchoolType = "Private")
dataset_AUS_Catholic
<- filter(dataset_AUS, as.numeric(STRATUM) == 1 |as.numeric(STRATUM) == 2 | as.numeric(STRATUM) == 6 | as.numeric(STRATUM) == 7 |as.numeric(STRATUM) == 12 | as.numeric(STRATUM) == 13 | as.numeric(STRATUM) == 18 |as.numeric(STRATUM) == 19 | as.numeric(STRATUM) == 24 | as.numeric(STRATUM) == 25 |as.numeric(STRATUM) == 30 | as.numeric(STRATUM) == 33 | as.numeric(STRATUM) == 37) %>%
 add_column(SchoolType = "Catholic")


#Make seperate data sets for countries
dataset_AUS <- bind_rows(dataset_AUS_Catholic, dataset_AUS_Gov, dataset_AUS_Private)
dataset_NL <- filter(dataset_All, CNT == "Netherlands")
#Remove datasets used for merging
```

```r
remove(dataset_AUS_Catholic,dataset_AUS_Private,dataset_AUS_Gov,dataset_AUS_ACT,dataset_AUS_NT,dataset_AUS_NSW,dataset_AUS_QLD,dataset_AUS_VIC,dataset_AUS_WA,dataset_AUS_TAS,dataset_AUS_SA)

dataset_NL <- dataset_NL %>%
 mutate(STRATUM = relevel(STRATUM, ref = "Netherlands - PRO or VMBO")) #Relevel Dutch stratum to Private Education
#Factorise
dataset_AUS$State <- as.factor(dataset_AUS$State)
dataset_AUS$SchoolType <- as.factor(dataset_AUS$SchoolType)
```

Differences between NL and AUS

This chunk evaluates if Dutch and Australian students are equal in terms of the predictors.

```r
#Measure p values between NL and AUS

#T-tests Key Xs participants
t.test(dataset_AUS$ambition, dataset_NL$ambition)
##
##  Welch Two Sample t-test
##
## data:  dataset_AUS$ambition and dataset_NL$ambition
## t = 52.72, df = 10338, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4274205 0.4604318
## sample estimates:
## mean of x mean of y
##  3.219928  2.776002
t.test(dataset_AUS$bullying, dataset_NL$bullying)
##
##  Welch Two Sample t-test
##
## data:  dataset_AUS$bullying and dataset_NL$bullying
## t = 37.512, df = 16565, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2485633 0.2759717
## sample estimates:
## mean of x mean of y
##  1.416717  1.154450
t.test(dataset_AUS$senseofbelonging, dataset_NL$senseofbelonging)
```

```
##
##  Welch Two Sample t-test
##
## data:  dataset_AUS$senseofbelonging and dataset_NL$senseofbelonging
## t = 23.142, df = 10114, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1883547 0.2232165
## sample estimates:
## mean of x mean of y
##  2.530511  2.324726
t.test(dataset_AUS$teacherengagement, dataset_NL$teacherengagement)
##
##  Welch Two Sample t-test
##
## data:  dataset_AUS$teacherengagement and dataset_NL$teacherengagement
## t = 10.842, df = 8128.9, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1177271 0.1696941
## sample estimates:
## mean of x mean of y
##  2.133189  1.989479
```

Centering predictors.

This chunks center the predictors, this means that the overall mean of every predictor is 0. Understanding output will be easier.

```
##Center X variables##

#Center X variables - All

dataset_All$ambition <- scale(dataset_All$ambition, center = TRUE, scale = FALSE)
dataset_All$bullying <- scale(dataset_All$bullying, center = TRUE, scale = FALSE)
dataset_All$senseofbelonging <- scale(dataset_All$senseofbelonging, center = TRUE, scale =FALSE)
dataset_All$teacherengagement <- scale(dataset_All$teacherengagement, center = TRUE, scale =FALSE)

#Center X variables - Australia

dataset_AUS$ambition <- scale(dataset_AUS$ambition, center = TRUE, scale = FALSE)
dataset_AUS$bullying <- scale(dataset_AUS$bullying, center = TRUE, scale = FALSE)
dataset_AUS$senseofbelonging <- scale(dataset_AUS$senseofbelonging, center = TRUE, scale =FALSE)
```

```r
dataset_AUS$teacherengagement <- scale(dataset_AUS$teacherengagement, center = TRUE, scale = FALSE)

#Center X variables - Netherlands

dataset_NL$ambition <- scale(dataset_NL$ambition, center = TRUE, scale = FALSE)
dataset_NL$bullying <- scale(dataset_NL$bullying, center = TRUE, scale = FALSE)
dataset_NL$senseofbelonging <- scale(dataset_NL$senseofbelonging, center = TRUE, scale =FALSE)
dataset_NL$teacherengagement <- scale(dataset_NL$teacherengagement, center = TRUE, scale =FALSE)
```

Principal Component Analysis

```r
####Principal Component Analyses####
princomp(Ambition) %>%
 summary(loadings = TRUE)
## Importance of components:
##                          Comp.1    Comp.2    Comp.3     Comp.4     Comp.5
## Standard deviation     1.3194338 0.5928082 0.5555212 0.50544052 0.39700808
## Proportion of Variance 0.6186551 0.1248826 0.1096667 0.09078487 0.05601084
## Cumulative Proportion  0.6186551 0.7435376 0.8532043 0.94398916 1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5
## ST119Q01NA  0.369  0.642  0.135  0.206  0.625
## ST119Q02NA  0.327  0.450         0.328 -0.764
## ST119Q03NA  0.539 -0.616  0.127  0.550  0.105
## ST119Q04NA  0.369        -0.908 -0.187
## ST119Q05NA  0.575         0.375 -0.715 -0.105
princomp(Belonging) %>%
 summary(loadings = TRUE)
## Importance of components:
##                          Comp.1    Comp.2     Comp.3     Comp.4
## Standard deviation     1.4292492 0.6815272 0.56030586 0.49172873
## Proportion of Variance 0.5951759 0.1353305 0.09147022 0.07044993
## Cumulative Proportion  0.5951759 0.7305064 0.82197660 0.89242652
##                          Comp.5     Comp.6
## Standard deviation     0.45158296 0.40655238
## Proportion of Variance 0.05941615 0.04815733
## Cumulative Proportion  0.95184267 1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## ST034Q01TA  0.496  0.348  0.119  0.689  0.378
```

```
## ST034Q02TA  0.341 -0.492  0.567 -0.186  0.198 -0.496
## ST034Q03TA  0.350 -0.577 -0.695  0.211
## ST034Q04TA  0.467  0.309 -0.285 -0.663  0.396
## ST034Q05TA  0.252 -0.351  0.298                      0.846
## ST034Q06TA  0.484  0.292  0.101        -0.804 -0.138
```
```r
princomp(Bullying) %>%
 summary(loadings = TRUE)
```
```
## Importance of components:
##                         Comp.1     Comp.2     Comp.3     Comp.4
## Standard deviation     1.4015573 0.5640229 0.54103378 0.46323636
## Proportion of Variance 0.6367775 0.1031239 0.09488876 0.06956187
## Cumulative Proportion  0.6367775 0.7399014 0.83479014 0.90435201
##                         Comp.5     Comp.6
## Standard deviation     0.40176085 0.36557886
## Proportion of Variance 0.05232403 0.04332396
## Cumulative Proportion  0.95667604 1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## ST038Q03NA  0.461  0.482         0.739
## ST038Q04NA  0.523  0.466 -0.325 -0.626 -0.107
## ST038Q05NA  0.370 -0.323 -0.121         0.581 -0.635
## ST038Q06NA  0.306 -0.469 -0.231  0.160 -0.750 -0.212
## ST038Q07NA  0.309 -0.452 -0.286         0.291  0.731
## ST038Q08NA  0.433 -0.146  0.863 -0.186
```
```r
princomp(TA) %>%
 summary(loadings = TRUE)
```
```
## Importance of components:
##                         Comp.1     Comp.2     Comp.3     Comp.4
## Standard deviation     1.5664578 0.6613714 0.58665538 0.53788956
## Proportion of Variance 0.6523892 0.1162948 0.09150303 0.07692288
## Cumulative Proportion  0.6523892 0.7686840 0.86018698 0.93710986
##                         Comp.5
## Standard deviation     0.48635859
## Proportion of Variance 0.06289014
## Cumulative Proportion  1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5
## ST118Q01NA  0.436  0.407                0.802
## ST118Q02NA  0.418  0.622  0.347 -0.153 -0.542
```

```
## ST118Q03NA  0.473        -0.647  0.543 -0.246
## ST118Q04NA  0.456 -0.371 -0.271 -0.761
## ST118Q05NA  0.451 -0.555  0.622  0.318
```
```r
princomp(Teacher) %>%
 summary(loadings = TRUE)
```
```
## Importance of components:
##                         Comp.1     Comp.2     Comp.3     Comp.4
## Standard deviation    1.7010584 0.52576242 0.45244377 0.41154735
## Proportion of Variance 0.7901866 0.07548667 0.05590111 0.04625203
## Cumulative Proportion  0.7901866 0.86567330 0.92157441 0.96782644
##                           Comp.5
## Standard deviation     0.34324478
## Proportion of Variance 0.03217356
## Cumulative Proportion  1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5
## ST104Q01NA  0.389  0.770  0.187  0.470
## ST104Q02NA  0.450  0.313 -0.219 -0.802
## ST104Q03NA  0.463 -0.237 -0.507  0.230 -0.647
## ST104Q04NA  0.467 -0.362 -0.215  0.267  0.730
## ST104Q05NA  0.461 -0.350  0.783 -0.114 -0.196
```

Descriptives of variables (per country)

```r
####Descriptives - All Countries - All variables####
dplyr::select(dataset_All, CNT, GENDER, testanxiety, ambition, bullying, senseofbelonging, teacherengagement) %>%
 describeBy(digits = 2) %>%
 dplyr::select(n, mean, sd)
```
```
##                      n mean   sd
## CNT*             18819 1.73 0.45
## GENDER*          18819 1.50 0.50
## testanxiety      18819 2.55 0.70
## ambition         18280 0.00 0.58
## bullying         18212 0.00 0.56
## senseofbelonging 18174 0.00 0.57
## teacherengagement 14709 0.00 0.76
```
```r
####Descriptives per country for all variables####
dplyr::select(dataset_All, CNT, GENDER, STRATUM, testanxiety, ambition, senseofbelonging, teacherengagement, bullying) %>%
 describeBy(group = dataset_All$CNT, mat=TRUE, digits = 2) %>%
 dplyr::select(group1,n,mean,sd)
```

```
##                              group1      n   mean     sd
## CNT*1              Netherlands  5124   1.00   0.00
## CNT*2                Australia 13695   2.00   0.00
## GENDER*1           Netherlands  5124   1.51   0.50
## GENDER*2             Australia 13695   1.50   0.50
## STRATUM*1          Netherlands  5124  41.48   0.50
## STRATUM*2            Australia 13695  18.93  10.15
## testanxiety1       Netherlands  5124   2.15   0.61
## testanxiety2         Australia 13695   2.70   0.67
## ambition1          Netherlands  4992  -0.32   0.48
## ambition2            Australia 13288   0.12   0.56
## senseofbelonging1  Netherlands  5015  -0.15   0.52
## senseofbelonging2    Australia 13159   0.06   0.58
## teacherengagement1 Netherlands  4049  -0.10   0.69
## teacherengagement2   Australia 10660   0.04   0.78
## bullying1          Netherlands  5034  -0.19   0.32
## bullying2            Australia 13178   0.07   0.61
```
#All descriptives for country, gender
```
dplyr::select(dataset_All, CNT, GENDER) %>%
 describe()
## .
##
##  2  Variables      18819  Observations
## --------------------------------------------------------------------------------
## CNT
##        n  missing distinct
##     18819        0        2
##
## Value        Netherlands    Australia
## Frequency           5124        13695
## Proportion         0.272        0.728
## --------------------------------------------------------------------------------
## GENDER
##        n  missing distinct
##     18819        0        2
##
## Value          Male Female
## Frequency      9403   9416
## Proportion      0.5    0.5
## --------------------------------------------------------------------------------
```

58

```r
#Dutch descriptives for country, gender, stratum
dplyr::select(dataset_NL, GENDER, STRATUM) %>%
 describe()
## .
##
##  2  Variables      5124  Observations
## --------------------------------------------------------------------------------
## GENDER
##        n  missing distinct
##     5124        0        2
##
## Value         Male Female
## Frequency     2524   2600
## Proportion   0.493  0.507
## --------------------------------------------------------------------------------
## STRATUM
##        n  missing distinct
##     5124        0        3
##
## Value              Netherlands - PRO or VMBO       Netherlands - HAVO or VWO
## Frequency                           2665                            2458
## Proportion                          0.52                            0.48
##
## Value        Netherlands - Private Education
## Frequency                                  1
## Proportion                              0.00
## --------------------------------------------------------------------------------
#Australian descriptives for country, gender, state
dplyr::select(dataset_AUS, GENDER, State) %>%
 describe()
## .
##
##  2  Variables     13695  Observations
## --------------------------------------------------------------------------------
## GENDER
##        n  missing distinct
##    13695        0        2
##
## Value         Male Female
## Frequency     6879   6816
## Proportion   0.502  0.498
```

```
## ---------------------------------------------------------------------------
## State
##         n  missing distinct
##     13695        0        8
##
## ACT (817, 0.060), New South Wales (3210, 0.234), Northern Territory (494,
## 0.036), Queensland (2777, 0.203), South Australia (1600, 0.117), Tasmania
## (929, 0.068), Victoria (2104, 0.154), Western Australia (1764, 0.129)
## ---------------------------------------------------------------------------
```

T tests and chi-square tests of demographics

#T-tests and ChiSq tests demographics, Y variable
t.test(dataset_AUS**$**testanxiety, dataset_NL**$**testanxiety) *#T-test TA between the Netherlands and Australia*

```
##
##  Welch Two Sample t-test
##
## data:  dataset_AUS$testanxiety and dataset_NL$testanxiety
## t = 53.983, df = 10025, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.5348086 0.5751117
## sample estimates:
## mean of x mean of y
##  2.704023  2.149063
```
chisq.test(x = dataset_All**$**CNT, y = dataset_All**$**GENDER) *#T-test between country and gender*
```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  dataset_All$CNT and dataset_All$GENDER
## X-squared = 1.3695, df = 1, p-value = 0.2419
```

Cronbach's Alphas of predictors and Test Anxiety

#Cronbach's Alphas
alpha(Ambition) *#Ambition Alpha*
```
##
## Reliability analysis
## Call: alpha(x = Ambition)
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd median_r
##        0.83      0.83    0.81       0.5   5 0.0019  3.1 0.58     0.48
##
```

```
##   lower alpha upper     95% confidence boundaries
## 0.82 0.83 0.83
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## ST119Q01NA    0.80      0.80    0.76     0.50 4.0  0.0023 0.0052  0.48
## ST119Q02NA    0.80      0.80    0.76     0.50 3.9  0.0023 0.0070  0.48
## ST119Q03NA    0.79      0.80    0.75     0.49 3.9  0.0025 0.0061  0.49
## ST119Q04NA    0.81      0.82    0.79     0.53 4.6  0.0022 0.0062  0.52
## ST119Q05NA    0.77      0.78    0.74     0.47 3.5  0.0027 0.0057  0.45
##
##  Item statistics
##              n raw.r std.r r.cor r.drop mean   sd
## ST119Q01NA 18537  0.75  0.77  0.70   0.62  3.3 0.68
## ST119Q02NA 18537  0.75  0.78  0.70   0.63  3.4 0.61
## ST119Q03NA 18537  0.80  0.78  0.71   0.65  3.0 0.85
## ST119Q04NA 18537  0.71  0.72  0.60   0.56  3.0 0.71
## ST119Q05NA 18537  0.84  0.82  0.77   0.71  2.8 0.87
##
## Non missing response frequency for each item
##              1    2    3    4 miss
## ST119Q01NA 0.01 0.09 0.50 0.40    0
## ST119Q02NA 0.01 0.04 0.47 0.48    0
## ST119Q03NA 0.04 0.24 0.38 0.35    0
## ST119Q04NA 0.02 0.20 0.56 0.22    0
## ST119Q05NA 0.07 0.32 0.39 0.22    0
alpha(Belonging) #Sense of Belonging Alpha
##
## Reliability analysis
## Call: alpha(x = Belonging)
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd median_r
##      0.85      0.85    0.84     0.49 5.7 0.0017  2.5 0.57     0.45
##
##  lower alpha upper     95% confidence boundaries
## 0.85 0.85 0.85
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## ST034Q01TA    0.81      0.82    0.80     0.47 4.5  0.0021 0.0074  0.44
## ST034Q02TA    0.83      0.83    0.81     0.49 4.9  0.0019 0.0140  0.44
```

```
## ST034Q03TA       0.84        0.84      0.83       0.51 5.3   0.0018 0.0136  0.46
## ST034Q04TA       0.82        0.82      0.80       0.48 4.6   0.0021 0.0076  0.45
## ST034Q05TA       0.84        0.84      0.82       0.51 5.1   0.0019 0.0121  0.45
## ST034Q06TA       0.81        0.81      0.79       0.46 4.3   0.0022 0.0064  0.44
##
##  Item statistics
##               n raw.r std.r r.cor r.drop mean   sd
## ST034Q01TA 18455  0.81  0.79  0.75   0.69  2.9 0.84
## ST034Q02TA 18455  0.72  0.74  0.67   0.60  2.0 0.71
## ST034Q03TA 18455  0.70  0.70  0.60   0.55  2.2 0.76
## ST034Q04TA 18455  0.80  0.78  0.73   0.68  2.9 0.81
## ST034Q05TA 18455  0.68  0.72  0.63   0.57  2.0 0.58
## ST034Q06TA 18455  0.83  0.81  0.78   0.72  2.8 0.81
##
## Non missing response frequency for each item
##               1    2    3    4    5 miss
## ST034Q01TA 0.00 0.34 0.46 0.14 0.06    0
## ST034Q02TA 0.19 0.61 0.16 0.04 0.00    0
## ST034Q03TA 0.15 0.59 0.20 0.06 0.00    0
## ST034Q04TA 0.00 0.32 0.49 0.14 0.05    0
## ST034Q05TA 0.14 0.75 0.09 0.03 0.00    0
## ST034Q06TA 0.00 0.42 0.43 0.10 0.05    0
alpha(Bullying) #Bullying Alpha
##
## Reliability analysis
## Call: alpha(x = Bullying)
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd median_r
##        0.88      0.88    0.87      0.55 7.3 0.0014  1.3 0.56     0.56
##
##  lower alpha upper     95% confidence boundaries
## 0.87 0.88 0.88
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## ST038Q03NA      0.85      0.86    0.84      0.55 6.1   0.0017 0.0033  0.56
## ST038Q04NA      0.85      0.85    0.83      0.54 5.8   0.0018 0.0039  0.57
## ST038Q05NA      0.85      0.85    0.83      0.53 5.6   0.0017 0.0045  0.51
## ST038Q06NA      0.86      0.87    0.85      0.56 6.4   0.0016 0.0042  0.57
## ST038Q07NA      0.86      0.86    0.84      0.56 6.2   0.0016 0.0037  0.57
## ST038Q08NA      0.86      0.87    0.85      0.56 6.4   0.0016 0.0046  0.58
```

```
##
##   Item statistics
##               n raw.r std.r r.cor r.drop mean   sd
## ST038Q03NA 18522  0.81  0.79  0.74   0.70  1.4 0.78
## ST038Q04NA 18522  0.83  0.81  0.77   0.72  1.5 0.85
## ST038Q05NA 18522  0.82  0.83  0.80   0.74  1.3 0.64
## ST038Q06NA 18522  0.74  0.76  0.69   0.64  1.2 0.61
## ST038Q07NA 18522  0.76  0.78  0.72   0.66  1.2 0.60
## ST038Q08NA 18522  0.78  0.76  0.69   0.66  1.4 0.78
##
## Non missing response frequency for each item
##              1    2    3    4 miss
## ST038Q03NA 0.72 0.18 0.06 0.04    0
## ST038Q04NA 0.66 0.21 0.07 0.05    0
## ST038Q05NA 0.83 0.11 0.03 0.03    0
## ST038Q06NA 0.83 0.12 0.03 0.02    0
## ST038Q07NA 0.86 0.09 0.03 0.02    0
## ST038Q08NA 0.72 0.18 0.06 0.04    0
```

```r
alpha(Teacher) #Teacher Engagement Alpha
```

```
##
## Reliability analysis
## Call: alpha(x = Teacher)
##
##   raw_alpha std.alpha G6(smc) average_r S/N     ase mean   sd median_r
##      0.93      0.93    0.92     0.73  14 0.00087  2.1 0.76     0.74
##
##  lower alpha upper     95% confidence boundaries
## 0.93 0.93 0.93
##
##   Reliability if an item is dropped:
##            raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## ST104Q01NA      0.93      0.93    0.91      0.77  13 0.00094 0.0015  0.76
## ST104Q02NA      0.91      0.91    0.90      0.73  11 0.00114 0.0055  0.71
## ST104Q03NA      0.91      0.91    0.89      0.72  10 0.00119 0.0025  0.73
## ST104Q04NA      0.91      0.91    0.89      0.72  10 0.00120 0.0018  0.73
## ST104Q05NA      0.92      0.92    0.90      0.74  11 0.00109 0.0040  0.74
##
##   Item statistics
##               n raw.r std.r r.cor r.drop mean   sd
## ST104Q01NA 14937  0.83  0.84  0.78   0.75  2.1 0.80
## ST104Q02NA 14937  0.90  0.90  0.86   0.83  2.0 0.86
```

```
## ST104Q03NA 14937  0.91  0.91  0.89   0.85  2.1 0.86
## ST104Q04NA 14937  0.91  0.91  0.89   0.86  2.1 0.87
## ST104Q05NA 14937  0.88  0.88  0.84   0.81  2.1 0.88
##
## Non missing response frequency for each item
##               1    2    3    4 miss
## ST104Q01NA 0.23 0.49 0.23 0.05    0
## ST104Q02NA 0.30 0.43 0.23 0.05    0
## ST104Q03NA 0.26 0.43 0.25 0.06    0
## ST104Q04NA 0.25 0.43 0.25 0.07    0
## ST104Q05NA 0.28 0.41 0.24 0.07    0
alpha(TA) #Test Anxiety Alpha
##
## Reliability analysis
## Call: alpha(x = TA)
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean  sd median_r
##      0.87      0.87    0.84      0.57 6.5 0.0015  2.6 0.7     0.56
##
##  lower alpha upper     95% confidence boundaries
## 0.86 0.87 0.87
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se   var.r med.r
## ST118Q01NA     0.83      0.83    0.79      0.55 4.9   0.0020 0.0025  0.55
## ST118Q02NA     0.85      0.85    0.81      0.58 5.5   0.0018 0.0012  0.58
## ST118Q03NA     0.83      0.83    0.80      0.55 5.0   0.0020 0.0036  0.55
## ST118Q04NA     0.84      0.84    0.80      0.56 5.1   0.0020 0.0034  0.55
## ST118Q05NA     0.85      0.85    0.81      0.58 5.5   0.0018 0.0029  0.59
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean   sd
## ST118Q01NA 18819  0.83  0.83  0.78   0.72  2.6 0.83
## ST118Q02NA 18819  0.78  0.79  0.71   0.66  2.7 0.85
## ST118Q03NA 18819  0.83  0.82  0.76   0.71  2.7 0.89
## ST118Q04NA 18819  0.81  0.81  0.75   0.70  2.3 0.87
## ST118Q05NA 18819  0.79  0.79  0.71   0.66  2.5 0.89
##
## Non missing response frequency for each item
##               1    2    3    4 miss
## ST118Q01NA 0.10 0.35 0.42 0.12    0
```

```
## ST118Q02NA 0.08 0.32 0.43 0.17    0
## ST118Q03NA 0.11 0.29 0.42 0.18    0
## ST118Q04NA 0.17 0.44 0.28 0.10    0
## ST118Q05NA 0.13 0.36 0.37 0.15    0
```

It seems that Sense of Belonging is not linear with TA (see further in the script). Sense of Belonging will be made a categorical variable (Low, for Sense of Belonging <0; Medium, for Sense of Belonging 0<x<1; and High, for Sense of Belonging >1).

```r
#Make Dummy variables for categories
dataset_All$SOB_dummy <- 0 #Make new variable SOB_dummy
dataset_All$SOB_dummy <- if_else(dataset_All$senseofbelonging<0,"Low", #if sense of belonging <0, give value "Low". Otherwise...
                          if_else(dataset_All$senseofbelonging>1,"High","Medium")) #...if sense of belonging >1, give value "High". Otherwise,
give value "Medium".
dataset_All$SOB_dummy <- as_factor(dataset_All$SOB_dummy, levels =c("Low","Medium","High"))


#Netherlands:

dataset_NL$SOB_dummy <- 0
dataset_NL$SOB_dummy <- if_else(dataset_NL$senseofbelonging<0,"Low", #if sense of belonging <0, give value "Low". Otherwise...
                          if_else(dataset_NL$senseofbelonging>1,"High","Medium")) #...if sense of belonging >1, give value "High". Otherwise,
give value "Medium".
dataset_NL$SOB_dummy <- as_factor(dataset_NL$SOB_dummy, levels =c("Low","Medium","High"))


#Australia

dataset_AUS$SOB_dummy <- 0
dataset_AUS$SOB_dummy <- if_else(dataset_AUS$senseofbelonging<0,"Low", #if sense of belonging <0, give value "Low". Otherwise...
                          if_else(dataset_AUS$senseofbelonging>1,"High","Medium")) #...if sense of belonging >1, give value "High". Otherwise,
give value "Medium".
dataset_AUS$SOB_dummy <- as_factor(dataset_AUS$SOB_dummy, levels =c("Low","Medium","High"))


describe(dataset_All$SOB_dummy) #General Sense of Belonging descriptives
describe(dataset_NL$SOB_dummy) #Dutch Sense of Belonging descriptives
describe(dataset_AUS$SOB_dummy) #Australian Sense of Belonging descriptives
```

Assumptions for the mediation analysis:

Linear Regression Models:

In here, a multiple regression model for the independent variables are made.

```r
#Model: Y ~ Xs + gender + country
```

```r
multreg_GOOD <- lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement +CNT + GENDER, data = dataset_All)

multregNoCNT <- lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement +GENDER, data = dataset_All)

multregOnlyCountry <- lm(testanxiety ~ CNT + GENDER, data = dataset_All)

#Model summary of good multiple regression
summary(multreg_GOOD)
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER, data = dataset_All)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.22964 -0.38219  0.01997  0.40250  2.25492
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.977892   0.012394 159.580   <2e-16 ***
## ambition          0.135485   0.009828  13.785   <2e-16 ***
## bullying          0.169197   0.010454  16.186   <2e-16 ***
## SOB_dummyMedium   0.210473   0.011310  18.609   <2e-16 ***
## SOB_dummyHigh     0.263605   0.022984  11.469   <2e-16 ***
## teacherengagement -0.014757   0.007002  -2.107   0.0351 *
## CNTAustralia      0.406443   0.012928  31.438   <2e-16 ***
## GENDERFemale      0.311811   0.010504  29.685   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6097 on 13835 degrees of freedom
##   (4976 observations deleted due to missingness)
## Multiple R-squared:  0.2357, Adjusted R-squared:  0.2353
## F-statistic: 609.5 on 7 and 13835 DF,  p-value: < 2.2e-16
#Assumption 1: linearity per X variable

#Plot for Ambition per country
ggplot(data = NULL, show.legend = TRUE) +
 geom_smooth(mapping = aes(x = dataset_AUS$ambition, y = dataset_AUS$testanxiety), color ="#00008B") +
 geom_smooth(mapping = aes(x = dataset_NL$ambition, y = dataset_NL$testanxiety), color ="Orange") +
```
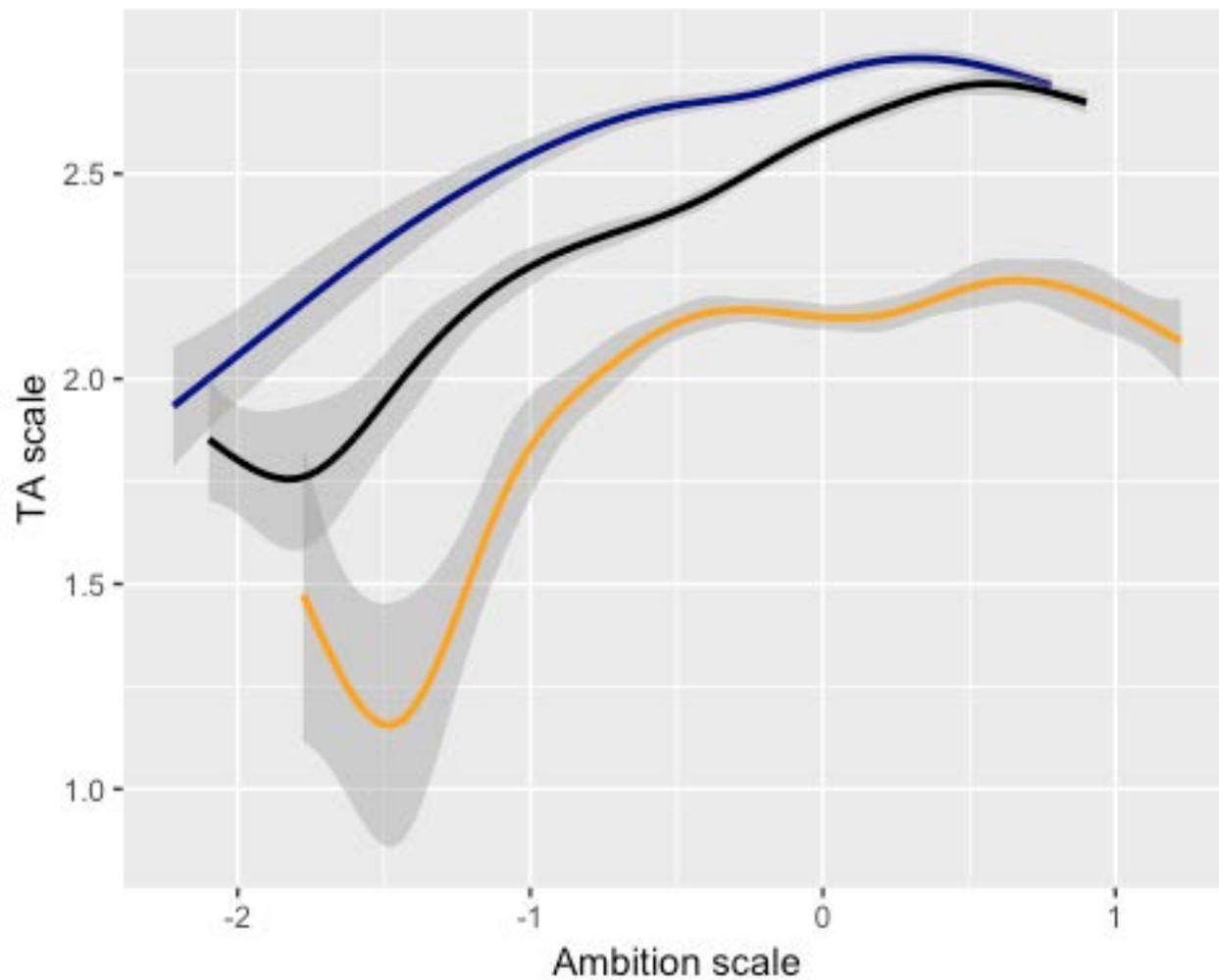
```
geom_smooth(mapping = aes(x = dataset_All$ambition, y = dataset_All$testanxiety), color ="Black") +
 labs(fill = "Country", name = "Ambition and TA per country", labels=c("Australia","The Netherlands"), x = "Ambition scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
#Skewness of Ambition
skewness(dataset_All$ambition, na.rm = TRUE)
## [1] -0.2177902
#Bullying graph
ggplot(data = NULL, show.legend = TRUE) +
 geom_smooth(mapping = aes(x = dataset_AUS$bullying, y = dataset_AUS$testanxiety), color ="#00008B") +
 geom_smooth(mapping = aes(x = dataset_NL$bullying, y = dataset_NL$testanxiety), color ="Orange") +
 geom_smooth(mapping = aes(x = dataset_All$bullying, y = dataset_All$testanxiety), color ="Black") +
 labs(name = "Bullying and TA per country", labels=c("Australia","The Netherlands"), x ="Bullying scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```
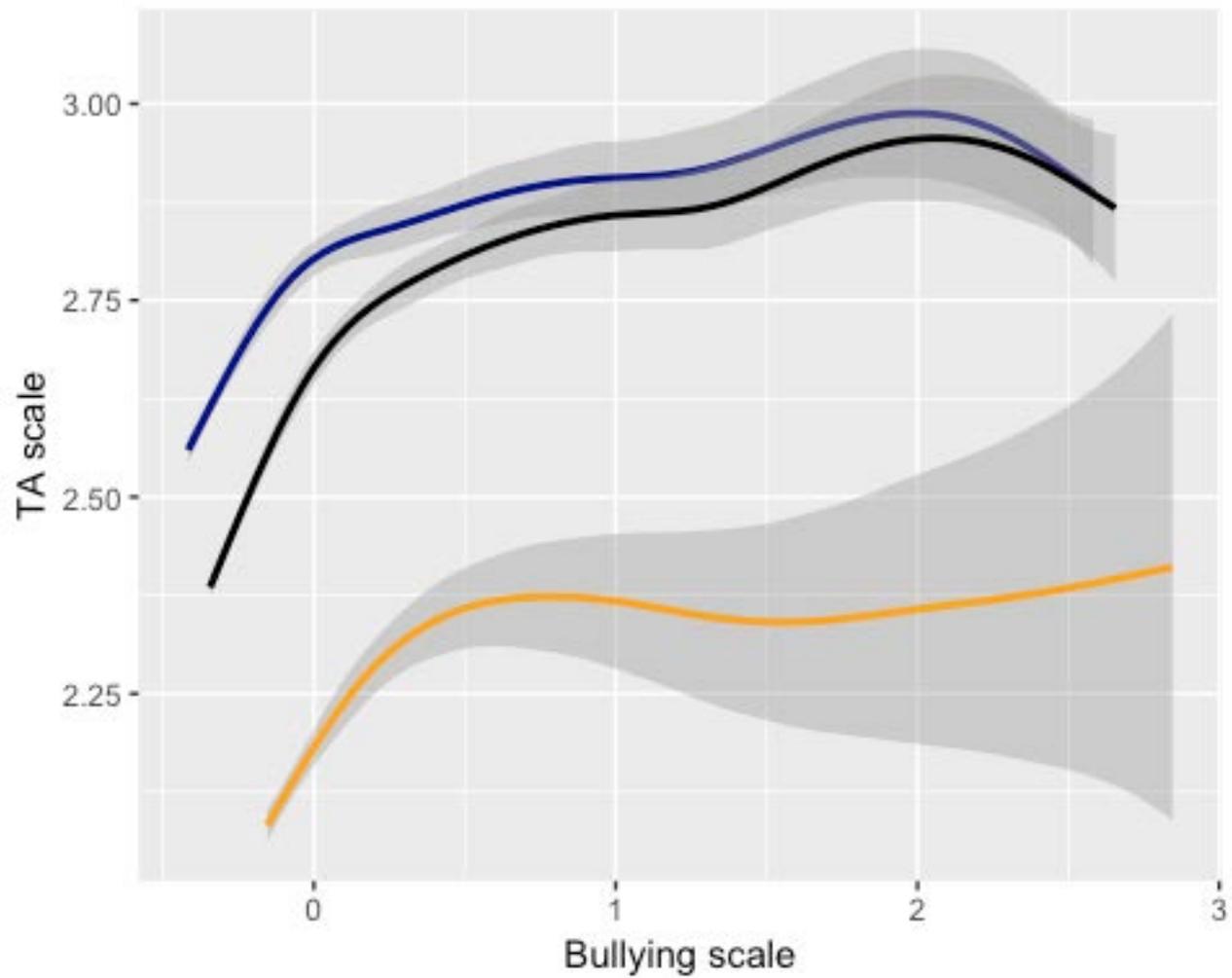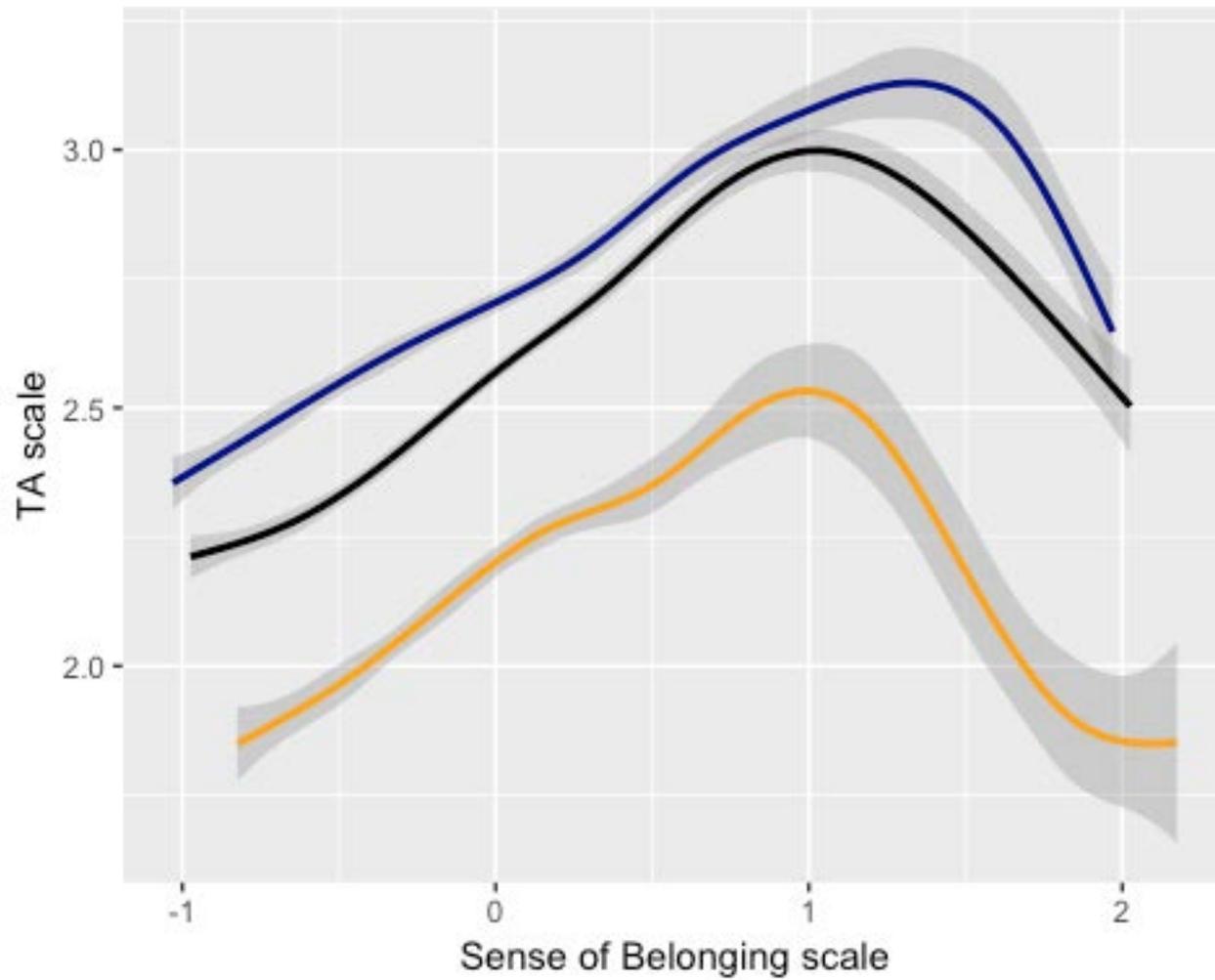
```
#Skewness of Bullying
skewness(dataset_All$bullying, na.rm = TRUE)
## [1] 2.429485
#Old Sense of Belonging in the classroom plot
ggplot(data = NULL, show.legend = TRUE) +
```

```r
geom_smooth(mapping = aes(x = dataset_AUS$senseofbelonging, y = dataset_AUS$testanxiety), color = "#00008B") +
geom_smooth(mapping = aes(x = dataset_NL$senseofbelonging, y = dataset_NL$testanxiety), color = "Orange") +
geom_smooth(mapping = aes(x = dataset_All$senseofbelonging, y = dataset_All$testanxiety), color = "Black") +
labs(name = "Sense of Belonging and TA per country", labels=c("Australia","The Netherlands"), x = "Sense of Belonging scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
#Skewness of Sense of Belonging
skewness(dataset_All$senseofbelonging, na.rm = TRUE)
## [1] 0.8554281
#SOB_Dummy interaction plot
#Interaction plot SOB x CNT
```

```r
interaction.plot(x.factor = dataset_All$SOB_dummy, trace.factor = dataset_All$CNT, response =dataset_All$testanxiety, xlab = "Sense of
Belonging", ylab = "Test Anxiety", trace.label ="Country")
```



```r
#Teacher Engagement regression plot
ggplot(data = NULL, show.legend = TRUE) +
 geom_smooth(mapping = aes(x = dataset_AUS$teacherengagement, y =dataset_AUS$testanxiety), color = "#00008B") +
```

```r
geom_smooth(mapping = aes(x = dataset_NL$teacherengagement, y = dataset_NL$testanxiety), color = "Orange") +
geom_smooth(mapping = aes(x = dataset_All$teacherengagement, y = dataset_All$testanxiety), color = "Black") +
labs(name = "Teacher engagement and TA per country", labels=c("Australia","The Netherlands"), x = "Teacher engagement scale", y = "TA scale")
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```
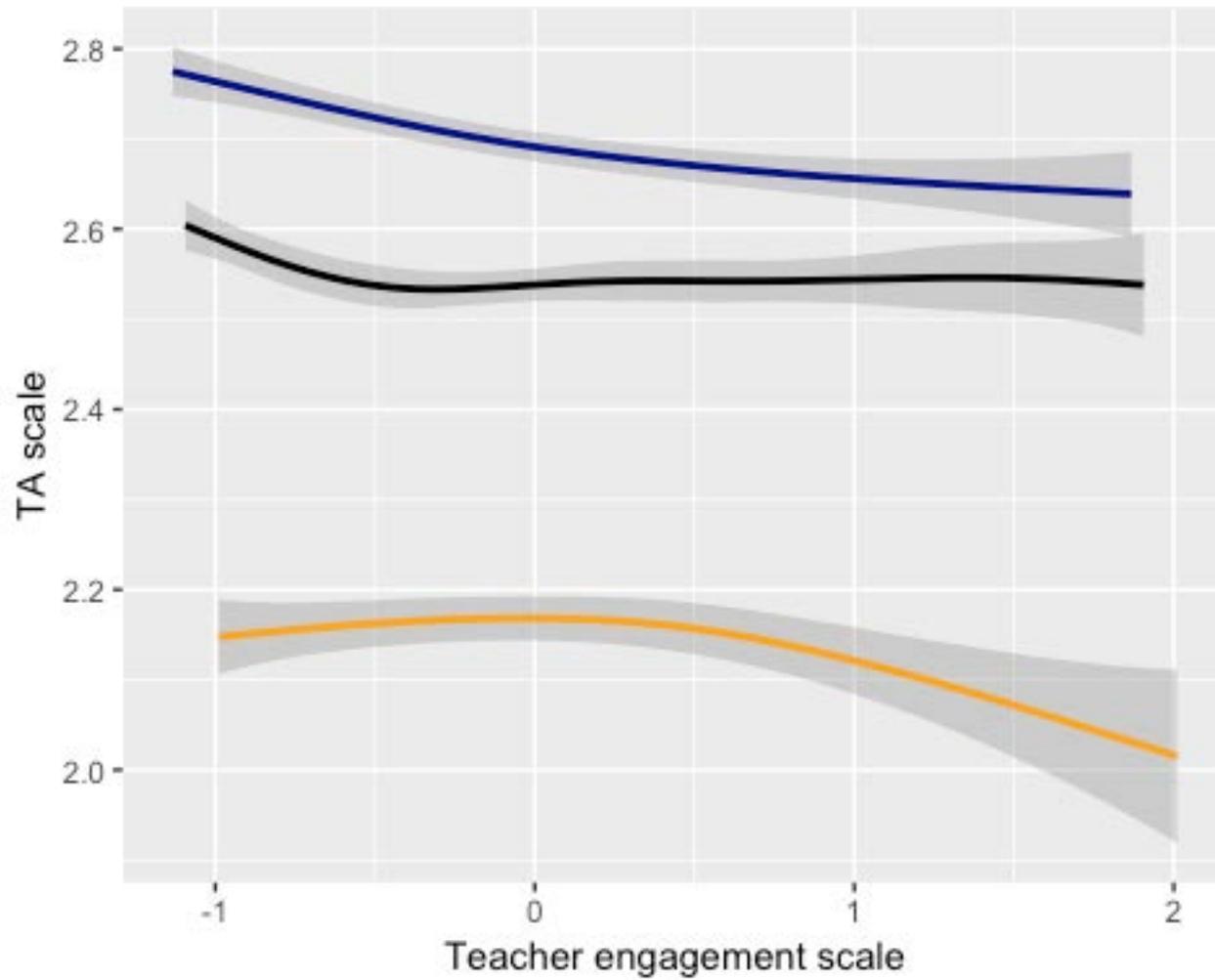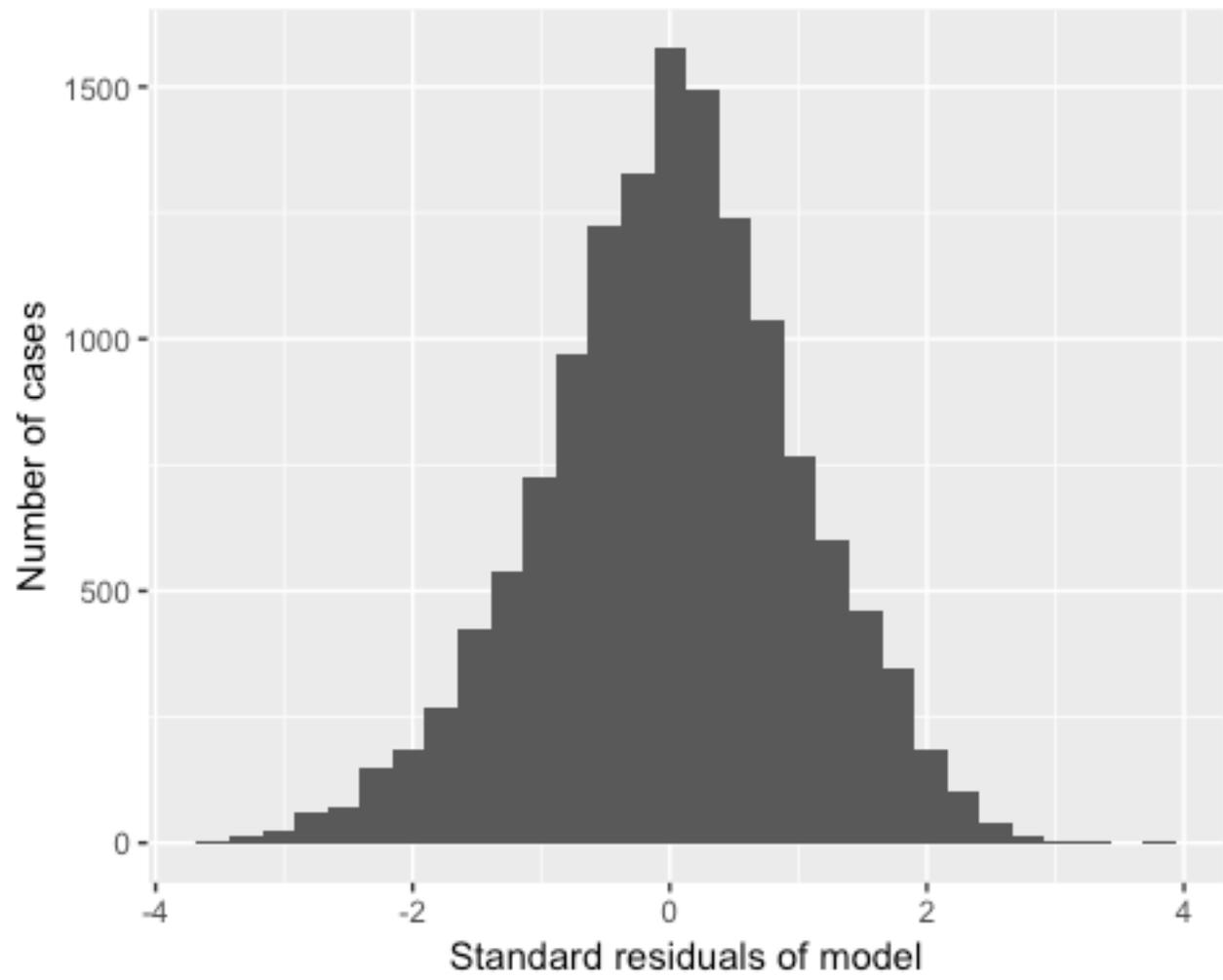
```
skewness(dataset_All$teacherengagement, na.rm = TRUE)
## [1] 0.4303407
#Assumption 2: Correlation Table between Key Xs

dplyr::select(dataset_All, ambition, bullying, senseofbelonging, teacherengagement) %>%
```
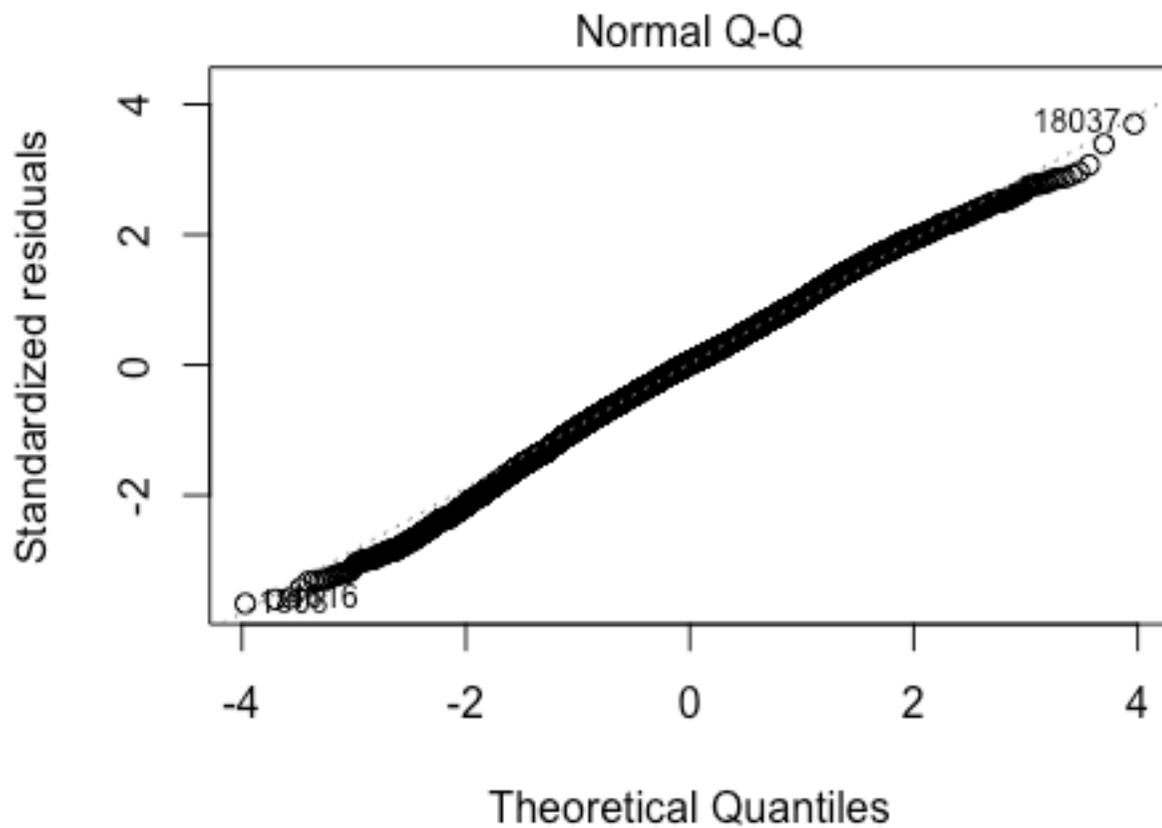
```r
as.matrix() %>%
rcorr(type = c("pearson"))
```
```
##                ambition bullying senseofbelonging teacherengagement
## ambition           1.00     0.04            -0.08              0.12
## bullying           0.04     1.00             0.38              0.01
## senseofbelonging  -0.08     0.38             1.00             -0.07
## teacherengagement  0.12     0.01            -0.07              1.00
##
## n
##                ambition bullying senseofbelonging teacherengagement
## ambition          18280    17743            17724             14390
## bullying          17743    18212            17741             14449
## senseofbelonging  17724    17741            18174             14357
## teacherengagement 14390    14449            14357             14709
##
## P
##                ambition bullying senseofbelonging teacherengagement
## ambition                 0.0000   0.0000           0.0000
## bullying        0.0000            0.0000           0.0932
## senseofbelonging 0.0000  0.0000                    0.0000
## teacherengagement 0.0000 0.0932   0.0000
```
```r
#Variance Influence Factor of multiple regression
vif(multreg_GOOD)
```
```
##           ambition          bullying    SOB_dummyMedium      SOB_dummyHigh
##             1.1994            1.1656             1.1873             1.1811
## teacherengagement        CNTAustralia       GENDERFemale
##             1.0412            1.2491             1.0272
```
```r
#Standard residuals of cases
augment(multreg_GOOD) %>%
 ggplot() +
 geom_histogram(mapping = aes(x = .std.resid)) +
 labs(name = "Standard residuals", y = "Number of cases", x = "Standard residuals of model")
```
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
#QQplot
plot(multreg_GOOD, 2)
```

Normal Q-Q

Standardized residuals

Theoretical Quantiles

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

```
#Residuals VS Fitted plot
plot(multreg_GOOD, 1)
```

Residuals vs Fitted

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

```
# Number of standard residuals >3 in multiple regression
augment(multreg_GOOD) %>%
 filter(.std.resid > 3 | .std.resid < -3) %>%
 dplyr::select(.std.resid)
```

```
## # A tibble: 28 x 1
##    .std.resid
##        <dbl>
##  1      -3.43
##  2      -3.02
##  3      -3.22
##  4      -3.30
##  5      -3.01
##  6      -3.17
##  7      -3.66
##  8      -3.25
##  9      -3.20
## 10      -3.02
## # … with 18 more rows
```
```r
#Cooks Distance
summary(augment(multreg_GOOD)$.cooksd)
```
```
##      Min.   1st Qu.   Median      Mean   3rd Qu.      Max.
## 0.000e+00 5.250e-06 2.488e-05 8.221e-05 8.483e-05 4.895e-03
```
```r
#Leverage plot
plot(multreg_GOOD, 5)
```

Residuals vs Leverage

stanxiety ~ ambition + bullying + SOB_dummy + teacherengagemen

Mediation Models

```
#Make dummy variable where Australian students are 1 and Dutch students are 0
dataset_All$isAustralian_dummy <- 0
dataset_All$isAustralian_dummy <- if_else(dataset_All$CNT=="Australia",1,0)
```

```r
#Make dummy variable where female students are 1 and male students are 0
dataset_All$isFemale_dummy <- 0
dataset_All$isFemale_dummy <- if_else(dataset_All$GENDER == "Female",1,0)

#Make dummy variable SOB numeric
dataset_All$SOB_dummy2 <- as.numeric(dataset_All$SOB_dummy)

#Model without mediator
lm(formula = testanxiety ~ isAustralian_dummy + isFemale_dummy, data = dataset_All) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ isAustralian_dummy + isFemale_dummy,
##     data = dataset_All)
##
## Residuals:
##      Min      1Q   Median      3Q     Max
## -1.87031 -0.38108  0.06074  0.46074  2.01892
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.98108    0.01002  197.74   <2e-16 ***
## isAustralian_dummy  0.55818    0.01038   53.78   <2e-16 ***
## isFemale_dummy      0.33105    0.00924   35.83   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6338 on 18816 degrees of freedom
## Multiple R-squared:  0.1804, Adjusted R-squared:  0.1803
## F-statistic:  2071 on 2 and 18816 DF,  p-value: < 2.2e-16
#Mediation Ambition
psych::mediate(testanxiety ~ isAustralian_dummy + isFemale_dummy + (ambition), data =dataset_All) %>%
 summary() #Output
```

Mediation
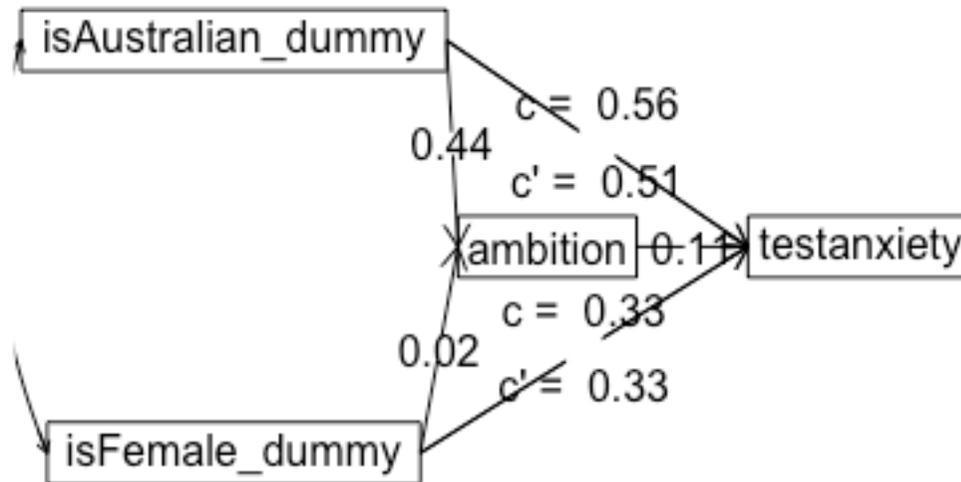

```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + isFemale_dummy +
##     (ambition), data = dataset_All)
##
##  Total effect estimates (c)
##                    testanxiety    se    t    df    Prob
```

```
## isAustralian_dummy        0.56 0.01 53.78 18816  0.00e+00
## isFemale_dummy            0.33 0.01 35.83 18816 5.44e-272
##
## Direct effect estimates      (c')
##                  testanxiety  se    t    df     Prob
## isAustralian_dummy        0.51 0.01 46.26 18815  0.0e+00
## isFemale_dummy            0.33 0.01 35.80 18815 1.3e-271
##
## R = 0.43 R2 = 0.19   F = 1450.27 on 3 and 18815 DF   p-value:  0
##
##  'a'  effect estimates
##                  ambition   se    t    df   Prob
## isAustralian_dummy    0.44 0.01 50.06 18816 0.0000
## isFemale_dummy        0.02 0.01  1.93 18816 0.0539
##
##  'b'  effect estimates
##          testanxiety  se    t    df     Prob
## ambition        0.11 0.01 13.07 18815 7.08e-39
##
##  'ab'  effect estimates
##                  testanxiety boot sd lower upper
## isAustralian_dummy        0.05 0.05  0  0.04  0.06
## isFemale_dummy            0.00 0.00  0  0.00  0.00
#Mediation Bullying
psych::mediate(testanxiety ~ isAustralian_dummy + isFemale_dummy + (bullying), data =dataset_All) %>%
 summary()
```
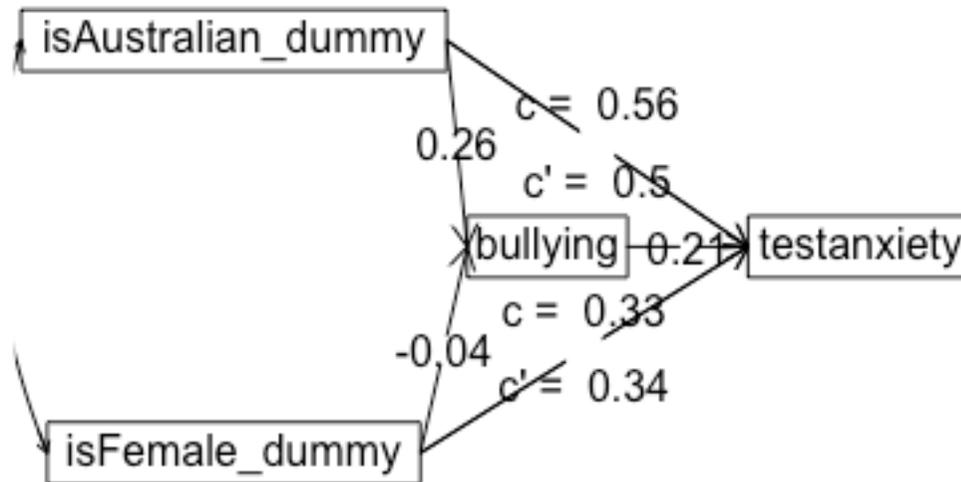
Mediation


```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + isFemale_dummy +
##      (bullying), data = dataset_All)
##
##  Total effect estimates (c)
##                     testanxiety   se     t    df     Prob
```

```
## isAustralian_dummy          0.56 0.01 53.78 18816  0.00e+00
## isFemale_dummy              0.33 0.01 35.83 18816 5.44e-272
## 
## Direct effect estimates      (c')
##                    testanxiety   se    t    df      Prob
## isAustralian_dummy          0.50 0.01 48.20 18815  0.00e+00
## isFemale_dummy              0.34 0.01 37.39 18815 3.67e-295
## 
## R = 0.45 R2 = 0.21   F = 1632.32 on 3 and 18815 DF   p-value:  0
## 
##  'a'  effect estimates
##                    bullying   se    t    df      Prob
## isAustralian_dummy     0.26 0.01 29.50 18816 5.71e-187
## isFemale_dummy        -0.04 0.01 -5.56 18816  2.76e-08
## 
##  'b'  effect estimates
##          testanxiety   se    t    df      Prob
## bullying        0.21 0.01 24.87 18815 2.37e-134
## 
##  'ab'  effect estimates
##                    testanxiety  boot sd lower upper
## isAustralian_dummy          0.05  0.05  0  0.05  0.06
## isFemale_dummy             -0.01 -0.01  0 -0.01 -0.01
```

```r
#Mediation Sense of Belonging
psych::mediate(testanxiety ~ isAustralian_dummy + isFemale_dummy + (SOB_dummy2), data =dataset_All) %>%
  summary()
```
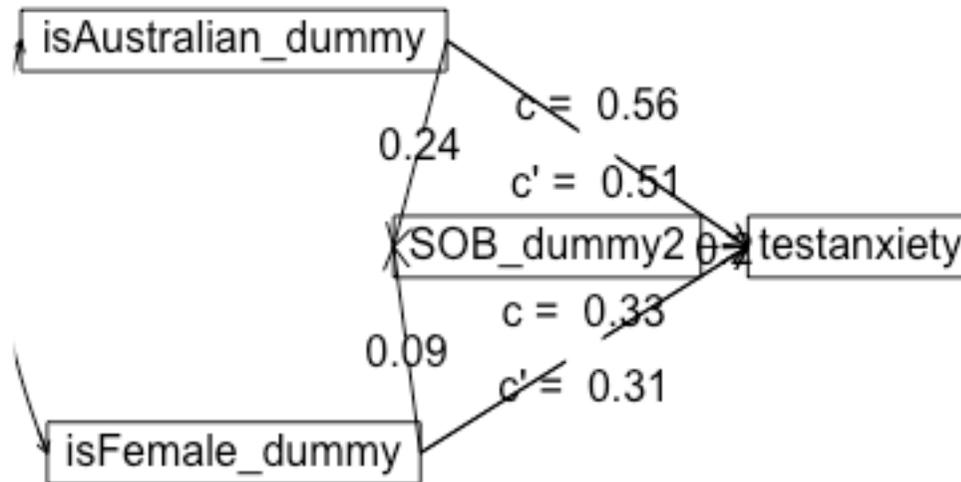
Mediation


```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + isFemale_dummy +
##     (SOB_dummy2), data = dataset_All)
##
## Total effect estimates (c)
##                  testanxiety   se    t    df    Prob
```

```
## isAustralian_dummy        0.56 0.01 53.78 18816  0.00e+00
## isFemale_dummy            0.33 0.01 35.83 18816 5.44e-272
##
## Direct effect estimates      (c')
##                   testanxiety   se    t    df    Prob
## isAustralian_dummy        0.51 0.01 49.26 18815  0e+00
## isFemale_dummy            0.31 0.01 34.36 18815 5e-251
##
## R = 0.46 R2 = 0.21   F = 1676.43 on 3 and 18815 DF   p-value:  0
##
##  'a'  effect estimates
##                   SOB_dummy2   se    t    df      Prob
## isAustralian_dummy        0.24 0.01 24.31 18816 1.61e-128
## isFemale_dummy            0.09 0.01 10.43 18816  2.16e-25
##
##  'b'  effect estimates
##          testanxiety   se    t    df   Prob
## SOB_dummy2        0.2 0.01 26.96 18815 4e-157
##
##  'ab'  effect estimates
##                   testanxiety boot sd lower upper
## isAustralian_dummy        0.05 0.05  0  0.04  0.05
## isFemale_dummy            0.02 0.02  0  0.01  0.02
```
#Mediation Teacher Engagement
```
psych::mediate(testanxiety ~ isAustralian_dummy + isFemale_dummy + (teacherengagement), data = dataset_All) %>%
  summary()
```
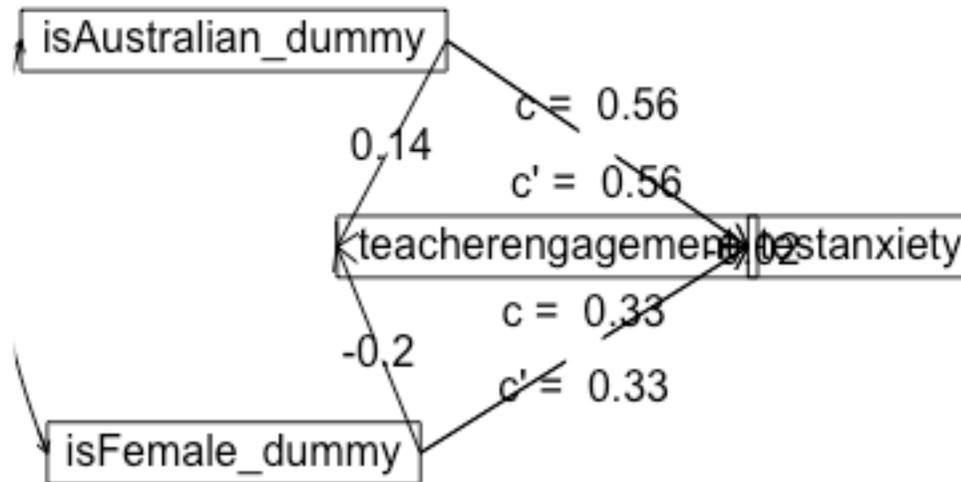
Mediation


```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + isFemale_dummy +
##     (teacherengagement), data = dataset_All)
##
##  Total effect estimates (c)
##                  testanxiety   se    t    df    Prob
```

```
## isAustralian_dummy        0.56 0.01 53.78 18816  0.00e+00
## isFemale_dummy            0.33 0.01 35.83 18816 5.44e-272
##
## Direct effect estimates     (c')
##                 testanxiety   se    t    df      Prob
## isAustralian_dummy        0.56 0.01 53.84 18815  0.00e+00
## isFemale_dummy            0.33 0.01 35.15 18815 2.87e-262
##
## R = 0.43 R2 = 0.18   F = 1384.12 on 3 and 18815 DF   p-value:  0
##
##  'a'  effect estimates
##                 teacherengagement   se     t    df     Prob
## isAustralian_dummy          0.14 0.01  11.64 18816 3.40e-31
## isFemale_dummy             -0.20 0.01 -18.03 18816 4.91e-72
##
##  'b'  effect estimates
##                 testanxiety   se    t    df    Prob
## teacherengagement      -0.02 0.01 -2.87 18815 0.00412
##
##  'ab'  effect estimates
##                 testanxiety boot sd lower upper
## isAustralian_dummy        0    0  0     0  0.00
## isFemale_dummy            0    0  0     0  0.01
#Only teacher engagement not significant in mediation model;
psych::mediate(testanxiety ~ isAustralian_dummy + isFemale_dummy + (ambition) + (bullying) +(senseofbelonging), data = dataset_All) %>%
  summary()
```

Mediation

isAustralian_dummy   ambition

0.26 c = 0.56

0.21   0.15

c' = 0.41

bullying   0.13   testanxiety

c = 0.33

0.02   0.22

0.04 c = 0.32

isFemale_dummy   senseofbelonging

```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + isFemale_dummy +
##     (ambition) + (bullying) + (senseofbelonging), data = dataset_All)
##
##  Total effect estimates (c)
##                     testanxiety    se     t    df     Prob
```
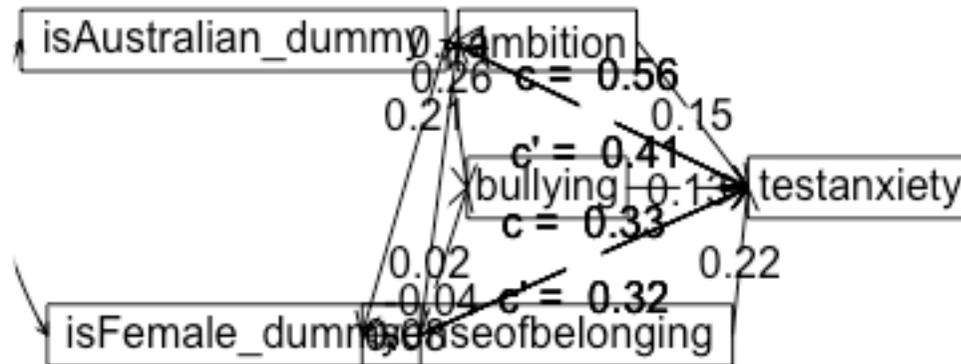
```
## isAustralian_dummy        0.56 0.01 53.78 18816  0.00e+00
## isFemale_dummy            0.33 0.01 35.83 18816 5.44e-272
##
## Direct effect estimates     (c')
##                      testanxiety   se     t    df      Prob
## isAustralian_dummy        0.41 0.01 37.50 18813 6.82e-297
## isFemale_dummy            0.32 0.01 35.44 18813 2.21e-266
##
## R = 0.49 R2 = 0.24   F = 1201.47 on 5 and 18813 DF   p-value:  0
##
##  'a'  effect estimates
##                      ambition   se     t    df    Prob
## isAustralian_dummy     0.44 0.01 50.06 18816 0.0000
## isFemale_dummy         0.02 0.01  1.93 18816 0.0539
##                      bullying   se     t    df      Prob
## isAustralian_dummy     0.26 0.01 29.50 18816 5.71e-187
## isFemale_dummy        -0.04 0.01 -5.56 18816  2.76e-08
##                      senseofbelonging   se     t    df      Prob
## isAustralian_dummy            0.21 0.01 22.59 18816 1.82e-111
## isFemale_dummy                0.08 0.01  9.85 18816  7.64e-23
##
##  'b'  effect estimates
##                      testanxiety   se     t    df      Prob
## ambition              0.15 0.01 17.97 18813  1.33e-71
## bullying              0.13 0.01 15.03 18813  8.63e-51
## senseofbelonging      0.22 0.01 25.97 18813 4.29e-146
##
##  'ab'  effect estimates
##                      testanxiety boot   sd lower upper
## isAustralian_dummy      0.15 0.15 0.01  0.14  0.16
## isFemale_dummy          0.01 0.01 0.00  0.01  0.02
##
##  'ab' effects estimates for each mediator
##                      ambition boot sd lower upper
## isAustralian_dummy     0.07 0.07  0  0.06  0.07
## isFemale_dummy         0.00 0.00  0  0.00  0.00
##                      bullying  boot sd lower upper
## isAustralian_dummy     0.03  0.03  0  0.03  0.04
## isFemale_dummy        -0.01 -0.01  0 -0.01  0.00
##                      senseofbelonging boot sd lower upper
```

```
## isAustralian_dummy             0.05 0.05  0  0.04  0.05
## isFemale_dummy                 0.02 0.02  0  0.01  0.02
#Gender and teacher engagement removed (not used in thesis)
psych::mediate(testanxiety ~ isAustralian_dummy + (ambition) + (bullying) + (senseofbelonging), data = dataset_All) %>%
 summary()
```
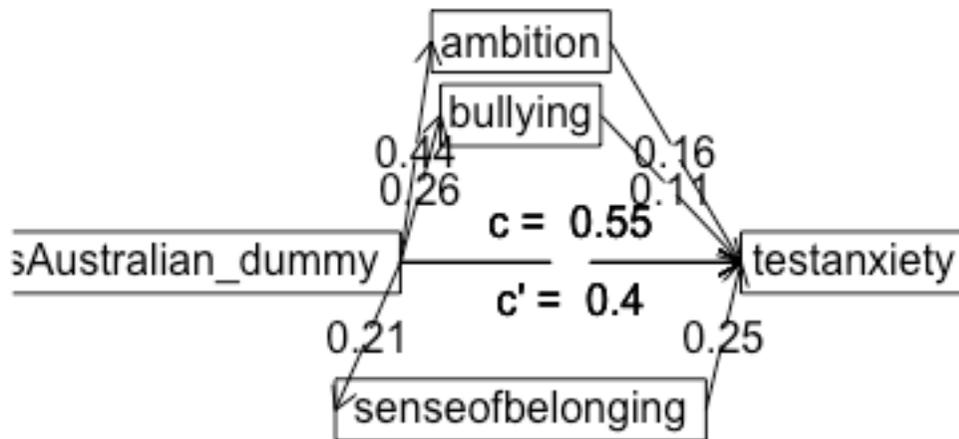
Mediation

```
## Call: psych::mediate(y = testanxiety ~ isAustralian_dummy + (ambition) +
##     (bullying) + (senseofbelonging), data = dataset_All)
##
##  Total effect estimates (c)
##                  testanxiety   se     t    df Prob
## isAustralian_dummy       0.55 0.01 51.74 18817    0
##
## Direct effect estimates     (c')
##                  testanxiety   se     t    df      Prob
## isAustralian_dummy       0.40 0.01 35.73 18814 1.44e-270
## ambition                 0.16 0.01 18.30 18814  3.69e-74
## bullying                 0.11 0.01 12.16 18814  6.87e-34
## senseofbelonging         0.25 0.01 28.51 18814 4.16e-175
##
## R = 0.44 R2 = 0.19   F = 1113.55 on 4 and 18814 DF   p-value:  0
##
##  'a'  effect estimates
##                  isAustralian_dummy   se     t    df      Prob
## ambition                       0.44 0.01 50.04 18817  0.00e+00
## bullying                       0.26 0.01 29.52 18817 2.76e-187
## senseofbelonging               0.21 0.01 22.44 18817 4.08e-110
##
##  'b'  effect estimates
##                  testanxiety   se     t    df      Prob
## ambition                 0.16 0.01 18.30 18814  3.69e-74
## bullying                 0.11 0.01 12.16 18814  6.87e-34
## senseofbelonging         0.25 0.01 28.51 18814 4.16e-175
##
##  'ab'  effect estimates
##                  testanxiety boot   sd lower upper
## isAustralian_dummy       0.15 0.15 0.01  0.14  0.16
##
##  'ab' effects estimates for each mediator
##                  ambition boot sd lower upper
## isAustralian_dummy    0.07 0.07  0  0.06  0.08
##                  bullying boot sd lower upper
## isAustralian_dummy    0.03 0.03  0  0.02  0.03
##                  senseofbelonging boot sd lower upper
## isAustralian_dummy            0.05 0.05  0  0.05  0.06
#### RQ1: Does 'Ambition' predict TA in both Australia and the Netherlands?####
```

```
AmbitionAllLM <- lm(dataset_All$testanxiety ~ dataset_All$ambition +  dataset_All$GENDER +dataset_All$CNT) %>%
 summary() #Summary of linear regression model of Ambition
AmbitionNlLM <- lm(dataset_NL$testanxiety ~ dataset_NL$ambition + dataset_NL$GENDER) %>%
 summary() #Summary of linear regression model of Ambition in the Netherlands
AmbitionAusLM <- lm(dataset_AUS$testanxiety ~ dataset_AUS$ambition +dataset_AUS$GENDER) %>%
 summary()#Summary of linear regression model of Ambition in Australia

#SOB multiple regression for all participants
lm(dataset_All$testanxiety ~ dataset_All$SOB_dummy + dataset_All$GENDER +dataset_All$CNT) %>%
 summary()
##
## Call:
## lm(formula = dataset_All$testanxiety ~ dataset_All$SOB_dummy +
##     dataset_All$GENDER + dataset_All$CNT)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.082 -0.401 -0.001  0.399  2.106
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   1.894167   0.010511  180.22   <2e-16 ***
## dataset_All$SOB_dummyMedium   0.226033   0.009720   23.25   <2e-16 ***
## dataset_All$SOB_dummyHigh     0.366953   0.019387   18.93   <2e-16 ***
## dataset_All$GENDERFemale      0.314003   0.009255   33.93   <2e-16 ***
## dataset_All$CNTAustralia      0.506837   0.010501   48.27   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6219 on 18169 degrees of freedom
##   (645 observations deleted due to missingness)
## Multiple R-squared:  0.2125, Adjusted R-squared:  0.2123
## F-statistic:  1225 on 4 and 18169 DF,  p-value: < 2.2e-16
#SOB multiple regression for Dutch participants
lm(dataset_NL$testanxiety ~ dataset_NL$SOB_dummy + dataset_NL$GENDER) %>%
 summary()
##
## Call:
## lm(formula = dataset_NL$testanxiety ~ dataset_NL$SOB_dummy +
##     dataset_NL$GENDER)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3935 -0.3478  0.0065  0.2994  2.0994
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.90064    0.01450 131.057  < 2e-16 ***
## dataset_NL$SOB_dummyMedium  0.24574    0.01700  14.459  < 2e-16 ***
## dataset_NL$SOB_dummyHigh    0.16816    0.03866   4.349 1.39e-05 ***
## dataset_NL$GENDERFemale     0.24712    0.01657  14.917  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5861 on 5011 degrees of freedom
##   (109 observations deleted due to missingness)
## Multiple R-squared:  0.08099,    Adjusted R-squared:  0.08044
## F-statistic: 147.2 on 3 and 5011 DF,  p-value: < 2.2e-16
```

```r
#SOB multiple regression for Australian participants
SOB_dummy_AusLM <- lm(dataset_AUS$testanxiety ~ dataset_AUS$SOB_dummy +dataset_AUS$GENDER) %>%
 summary()


#Bullying multiple regression for all participants
lm(dataset_All$testanxiety ~ dataset_All$bullying + dataset_All$GENDER + dataset_All$CNT) %>%
 summary()
```

```
##
## Call:
## lm(formula = dataset_All$testanxiety ~ dataset_All$bullying +
##     dataset_All$GENDER + dataset_All$CNT)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.30051 -0.38375  0.01625  0.41625  2.05423
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.016917   0.010031  201.08   <2e-16 ***
## dataset_All$bullying     0.206702   0.008422   24.54   <2e-16 ***
## dataset_All$GENDERFemale 0.335782   0.009226   36.40   <2e-16 ***
## dataset_All$CNTAustralia 0.502204   0.010539   47.65   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Residual standard error: 0.622 on 18208 degrees of freedom
##   (607 observations deleted due to missingness)
## Multiple R-squared:  0.2067, Adjusted R-squared:  0.2066
## F-statistic:  1582 on 3 and 18208 DF,  p-value: < 2.2e-16
```

#Bullying multiple regression for Dutch participants
```r
lm(dataset_NL$testanxiety ~ dataset_NL$bullying + dataset_NL$GENDER) %>%
 summary()
```
```
## 
## Call:
## lm(formula = dataset_NL$testanxiety ~ dataset_NL$bullying + dataset_NL$GENDER)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -1.75062 -0.38004  0.01996  0.36717  2.01996 
## 
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)    
## (Intercept)               2.01971    0.01187 170.083   <2e-16 ***
## dataset_NL$bullying       0.25686    0.02604   9.863   <2e-16 ***
## dataset_NL$GENDERFemale   0.25279    0.01669  15.146   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.5919 on 5031 degrees of freedom
##   (90 observations deleted due to missingness)
## Multiple R-squared:  0.05974,    Adjusted R-squared:  0.05936
## F-statistic: 159.8 on 2 and 5031 DF,  p-value: < 2.2e-16
```

#Bullying multiple regression for Australian participants
```r
lm(dataset_AUS$testanxiety ~ dataset_AUS$bullying + dataset_AUS$GENDER) %>%
 summary()
```
```
## 
## Call:
## lm(formula = dataset_AUS$testanxiety ~ dataset_AUS$bullying +
##     dataset_AUS$GENDER)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -2.30728 -0.40256 -0.00149  0.39851  1.56604 
## 
## Coefficients:
```

```
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.518272   0.007784  323.53   <2e-16 ***
## dataset_AUS$bullying     0.202316   0.009000   22.48   <2e-16 ***
## dataset_AUS$GENDERFemale 0.367530   0.011028   33.33   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6323 on 13175 degrees of freedom
##   (517 observations deleted due to missingness)
## Multiple R-squared:  0.1053, Adjusted R-squared:  0.1052
## F-statistic: 775.4 on 2 and 13175 DF,  p-value: < 2.2e-16
#Teacher engagement multiple regression for all participants

lm(dataset_All$testanxiety ~ dataset_All$teacherengagement + dataset_All$GENDER +dataset_All$CNT) %>%
 summary()
##
## Call:
## lm(formula = dataset_All$testanxiety ~ dataset_All$teacherengagement +
##     dataset_All$GENDER + dataset_All$CNT)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.88171 -0.39472  0.03001  0.43354  2.04767
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    1.986002   0.011235 176.775   <2e-16 ***
## dataset_All$teacherengagement -0.017662   0.006949  -2.542    0.011 *
## dataset_All$GENDERFemale       0.323429   0.010502  30.797   <2e-16 ***
## dataset_All$CNTAustralia       0.552961   0.011697  47.275   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6313 on 14705 degrees of freedom
##   (4110 observations deleted due to missingness)
## Multiple R-squared:  0.1798, Adjusted R-squared:  0.1796
## F-statistic:  1074 on 3 and 14705 DF,  p-value: < 2.2e-16
#Teacher engagement multiple regression for Dutch participants

lm(dataset_NL$testanxiety ~ dataset_NL$teacherengagement + dataset_NL$GENDER) %>%
 summary()
```

```
## 
## Call:
## lm(formula = dataset_NL$testanxiety ~ dataset_NL$teacherengagement +
##     dataset_NL$GENDER)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2706 -0.2706 -0.0353  0.3601  1.9762
## 
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.035364   0.013338 152.599   <2e-16 ***
## dataset_NL$teacherengagement  -0.005767   0.013546  -0.426     0.67
## dataset_NL$GENDERFemale        0.229567   0.018801  12.210   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.5945 on 4046 degrees of freedom
##   (1075 observations deleted due to missingness)
## Multiple R-squared:  0.03627,    Adjusted R-squared:  0.03579
## F-statistic: 76.13 on 2 and 4046 DF,  p-value: < 2.2e-16
```
#Teacher engagement multiple regression for Australian participants
`lm(dataset_AUS$testanxiety ~ dataset_AUS$teacherengagement + dataset_AUS$GENDER) %>%`
` summary()`
```
## 
## Call:
## lm(formula = dataset_AUS$testanxiety ~ dataset_AUS$teacherengagement +
##     dataset_AUS$GENDER)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9024 -0.4658  0.0690  0.4771  1.5178
## 
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.520163   0.008917 282.634   <2e-16 ***
## dataset_AUS$teacherengagement -0.020330   0.008096  -2.511    0.012 *
## dataset_AUS$GENDERFemale       0.359197   0.012594  28.522   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
```

```
## Residual standard error: 0.6438 on 10657 degrees of freedom
##   (3035 observations deleted due to missingness)
## Multiple R-squared:  0.07435,    Adjusted R-squared:  0.07418
## F-statistic:   428 on 2 and 10657 DF,  p-value: < 2.2e-16
```

Full multiple regression models

Research Question 6

RQ6: Does the model differ between Australia and the Netherlands?

#Interaction model Ambition x CNT
```r
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + CNT + GENDER +ambition:CNT, data = dataset_All) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER + ambition:CNT, data = dataset_All)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -2.23165 -0.38285  0.01962  0.40249  2.24784
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.975562   0.013675 144.470  < 2e-16 ***
## ambition             0.128224   0.020515   6.250 4.22e-10 ***
## bullying             0.169262   0.010455  16.189  < 2e-16 ***
## SOB_dummyMedium      0.210655   0.011320  18.609  < 2e-16 ***
## SOB_dummyHigh        0.263799   0.022990  11.474  < 2e-16 ***
## teacherengagement   -0.014810   0.007004  -2.115   0.0345 *
## CNTAustralia         0.408461   0.013864  29.462  < 2e-16 ***
## GENDERFemale         0.311571   0.010521  29.614  < 2e-16 ***
## ambition:CNTAustralia 0.009392  0.023293   0.403   0.6868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6097 on 13834 degrees of freedom
##   (4976 observations deleted due to missingness)
## Multiple R-squared:  0.2357, Adjusted R-squared:  0.2353
## F-statistic: 533.3 on 8 and 13834 DF,  p-value: < 2.2e-16
```

```r
#Interaction model Bullying x CNT
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + CNT + GENDER +bullying:CNT, data = dataset_All) %>%
  summary()
```

```
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER + bullying:CNT, data = dataset_All)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.22384 -0.38288  0.02046  0.40273  2.25795
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.982044   0.013609 145.646  < 2e-16 ***
## ambition              0.135214   0.009835  13.748  < 2e-16 ***
## bullying              0.191085   0.031409   6.084  1.2e-09 ***
## SOB_dummyMedium       0.209963   0.011332  18.529  < 2e-16 ***
## SOB_dummyHigh         0.264195   0.022999  11.487  < 2e-16 ***
## teacherengagement    -0.014826   0.007003  -2.117   0.0343 *
## CNTAustralia          0.402650   0.013910  28.947  < 2e-16 ***
## GENDERFemale          0.311897   0.010505  29.691  < 2e-16 ***
## bullying:CNTAustralia -0.024357   0.032958  -0.739   0.4599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6097 on 13834 degrees of freedom
##   (4976 observations deleted due to missingness)
## Multiple R-squared:  0.2357, Adjusted R-squared:  0.2353
## F-statistic: 533.4 on 8 and 13834 DF,  p-value: < 2.2e-16
```

```r
#Interaction model Sense of Belonnging x CNT
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + CNT + GENDER +SOB_dummy:CNT, data = dataset_All) %>%
  summary()
```

```
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER + SOB_dummy:CNT, data = dataset_All)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2728 -0.3826  0.0197  0.4010  2.2493
```

```
## 
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.982177   0.014218 139.413  < 2e-16 ***
## ambition                    0.136180   0.009835  13.846  < 2e-16 ***
## bullying                    0.162635   0.010526  15.451  < 2e-16 ***
## SOB_dummyMedium             0.228319   0.020842  10.954  < 2e-16 ***
## SOB_dummyHigh               0.027273   0.048562   0.562   0.5744
## teacherengagement          -0.014344   0.006996  -2.050   0.0403 *
## CNTAustralia                0.399665   0.016896  23.655  < 2e-16 ***
## GENDERFemale                0.309345   0.010507  29.441  < 2e-16 ***
## SOB_dummyMedium:CNTAustralia -0.018863  0.024431  -0.772   0.4401
## SOB_dummyHigh:CNTAustralia   0.299046   0.054920   5.445 5.27e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.609 on 13833 degrees of freedom
##   (4976 observations deleted due to missingness)
## Multiple R-squared:  0.2375, Adjusted R-squared:  0.237
## F-statistic: 478.8 on 9 and 13833 DF,  p-value: < 2.2e-16
#Interaction model Teacher Engagement x CNT
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + CNT + GENDER +teacherengagement:CNT, data = dataset_All) %>%
 summary()
## 
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + CNT + GENDER + teacherengagement:CNT,
##     data = dataset_All)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.23328 -0.38169  0.01927  0.40302  2.25483
## 
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.979862   0.012473 158.726   <2e-16 ***
## ambition                    0.135693   0.009829  13.805   <2e-16 ***
## bullying                    0.169127   0.010453  16.179   <2e-16 ***
## SOB_dummyMedium             0.210290   0.011311  18.592   <2e-16 ***
## SOB_dummyHigh               0.263182   0.022986  11.450   <2e-16 ***
## teacherengagement           0.002668   0.014272   0.187    0.852
```

```
## CNTAustralia                 0.404810   0.012980  31.187   <2e-16 ***
## GENDERFemale                  0.311736   0.010504  29.679   <2e-16 ***
## teacherengagement:CNTAustralia -0.022793  0.016268  -1.401    0.161
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6097 on 13834 degrees of freedom
##   (4976 observations deleted due to missingness)
## Multiple R-squared:  0.2358, Adjusted R-squared:  0.2354
## F-statistic: 533.6 on 8 and 13834 DF,  p-value: < 2.2e-16
#Question - What are the full models for different countries?

#Dutch multiple regression model
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER, data =dataset_NL) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER, data = dataset_NL)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6793 -0.3427  0.0238  0.3345  2.2332
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.918849   0.016616 115.485  < 2e-16 ***
## ambition          0.124803   0.019554   6.383 1.95e-10 ***
## bullying          0.195986   0.030324   6.463 1.15e-10 ***
## SOB_dummyMedium   0.223774   0.019592  11.422  < 2e-16 ***
## SOB_dummyHigh     0.093756   0.044025   2.130   0.0333 *
## teacherengagement -0.001097  0.013601  -0.081   0.9357
## GENDERFemale      0.235591   0.018782  12.543  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5767 on 3840 degrees of freedom
##   (1277 observations deleted due to missingness)
## Multiple R-squared:  0.09162,   Adjusted R-squared:  0.0902
## F-statistic: 64.55 on 6 and 3840 DF,  p-value: < 2.2e-16
```

```
#Australian full multiple regression model
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER, data =dataset_AUS) %>%
  summary()
## 
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER, data = dataset_AUS)
## 
## Residuals:
##      Min      1Q   Median      3Q      Max
## -2.18805 -0.39036  0.01667  0.41589  1.76251
## 
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.442277   0.010072 242.484  < 2e-16 ***
## ambition           0.129632   0.011399  11.373  < 2e-16 ***
## bullying           0.161082   0.011469  14.045  < 2e-16 ***
## SOB_dummyMedium    0.209483   0.014218  14.733  < 2e-16 ***
## SOB_dummyHigh      0.231679   0.030855   7.509 6.49e-14 ***
## teacherengagement -0.020417   0.008168  -2.500   0.0125 *
## GENDERFemale       0.344303   0.012656  27.205  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.6217 on 9989 degrees of freedom
##   (3699 observations deleted due to missingness)
## Multiple R-squared:  0.1379, Adjusted R-squared:  0.1373
## F-statistic: 266.2 on 6 and 9989 DF,  p-value: < 2.2e-16
#Question - Does TA in various Australian states or school types differ?

#Multiple regression in Australia with differences in states
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER + State, data = dataset_AUS) %>%
  summary()
## 
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER + State, data = dataset_AUS)
## 
## Residuals:
##      Min      1Q   Median      3Q      Max
## -2.17151 -0.39031  0.01765  0.41638  1.73067
```

```
## 
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.473173   0.028173  87.785  < 2e-16 ***
## ambition                 0.129141   0.011434  11.295  < 2e-16 ***
## bullying                 0.160532   0.011488  13.974  < 2e-16 ***
## SOB_dummyMedium          0.208861   0.014232  14.675  < 2e-16 ***
## SOB_dummyHigh            0.231107   0.030861   7.489 7.55e-14 ***
## teacherengagement       -0.019624   0.008194  -2.395   0.0166 *
## GENDERFemale             0.345001   0.012653  27.266  < 2e-16 ***
## StateNew South Wales    -0.016449   0.029781  -0.552   0.5807
## StateNorthern Territory -0.024369   0.043857  -0.556   0.5785
## StateQueensland         -0.002420   0.031178  -0.078   0.9381
## StateSouth Australia    -0.032250   0.032258  -1.000   0.3174
## StateTasmania           -0.082321   0.035677  -2.307   0.0211 *
## StateVictoria           -0.061771   0.031142  -1.984   0.0473 *
## StateWestern Australia  -0.040355   0.031945  -1.263   0.2065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.6214 on 9982 degrees of freedom
##   (3699 observations deleted due to missingness)
## Multiple R-squared:  0.1392, Adjusted R-squared:  0.1381
## F-statistic: 124.2 on 13 and 9982 DF,  p-value: < 2.2e-16
#Multiple regression in Australia with school types
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER +SchoolType, data = dataset_AUS) %>%
  summary()
## 
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER + SchoolType, data = dataset_AUS)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.18561 -0.38566  0.01889  0.41558  1.73496
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.470050   0.014800 166.890  < 2e-16 ***
## ambition     0.129206   0.011419  11.315  < 2e-16 ***
## bullying     0.160981   0.011491  14.009  < 2e-16 ***
```

```
## SOB_dummyMedium      0.210865    0.014247   14.800   < 2e-16 ***
## SOB_dummyHigh         0.231577    0.030863    7.503 6.75e-14 ***
## teacherengagement   -0.019887    0.008176   -2.432   0.01502 *
## GENDERFemale          0.344053    0.012652   27.193   < 2e-16 ***
## SchoolTypePrivate    -0.054518    0.018695   -2.916   0.00355 **
## SchoolTypePublic     -0.030512    0.015297   -1.995   0.04612 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6214 on 9987 degrees of freedom
##   (3699 observations deleted due to missingness)
## Multiple R-squared:  0.1386, Adjusted R-squared:  0.1379
## F-statistic: 200.9 on 8 and 9987 DF,  p-value: < 2.2e-16
#Multiple regression in the Netherlands with school type
lm(testanxiety ~ ambition + bullying + SOB_dummy + teacherengagement + GENDER +STRATUM, data = dataset_NL) %>%
 summary()
##
## Call:
## lm(formula = testanxiety ~ ambition + bullying + SOB_dummy +
##     teacherengagement + GENDER + STRATUM, data = dataset_NL)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7334 -0.3568  0.0129  0.3446  2.1906
##
## Coefficients:
##                                  Estimate Std. Error t value
## (Intercept)                       1.88185    0.01921  97.979
## ambition                          0.11656    0.01964   5.934
## bullying                          0.20239    0.03032   6.675
## SOB_dummyMedium                   0.22521    0.01957  11.510
## SOB_dummyHigh                     0.10138    0.04400   2.304
## teacherengagement                 0.00792    0.01378   0.575
## GENDERFemale                      0.23532    0.01876  12.547
## STRATUMNetherlands - HAVO or VWO  0.07248    0.01900   3.815
## STRATUMNetherlands - Private Education  0.17775    0.57633   0.308
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## ambition                         3.22e-09 ***
## bullying                         2.83e-11 ***
## SOB_dummyMedium                   < 2e-16 ***
```

```
## SOB_dummyHigh                         0.021261 *
## teacherengagement                     0.565636
## GENDERFemale                           < 2e-16 ***
## STRATUMNetherlands - HAVO or VWO       0.000139 ***
## STRATUMNetherlands - Private Education 0.757775
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5758 on 3838 degrees of freedom
##   (1277 observations deleted due to missingness)
## Multiple R-squared:  0.09506,    Adjusted R-squared:  0.09318
## F-statistic:  50.4 on 8 and 3838 DF,  p-value: < 2.2e-16
```

Citations

#Citations

```
citation()
##
## To cite R in publications use:
##
##   R Core Team (2018). R: A language and environment for
##   statistical computing. R Foundation for Statistical Computing,
##   Vienna, Austria. URL https://www.R-project.org/.
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {R: A Language and Environment for Statistical Computing},
##     author = {{R Core Team}},
##     organization = {R Foundation for Statistical Computing},
##     address = {Vienna, Austria},
##     year = {2018},
##     url = {https://www.R-project.org/},
##   }
##
## We have invested a lot of time and effort in creating R, please
## cite it when using it for data analysis. See also
## 'citation("pkgname")' for citing R packages.
citation(package = "tidyverse")
##
## To cite package 'tidyverse' in publications use:
```

```
##
##   Hadley Wickham (2017). tidyverse: Easily Install and Load the
##   'Tidyverse'. R package version 1.2.1.
##   https://CRAN.R-project.org/package=tidyverse
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {tidyverse: Easily Install and Load the 'Tidyverse'},
##     author = {Hadley Wickham},
##     year = {2017},
##     note = {R package version 1.2.1},
##     url = {https://CRAN.R-project.org/package=tidyverse},
##   }
citation(package = "haven")
##
## To cite package 'haven' in publications use:
##
##   Hadley Wickham and Evan Miller (2019). haven: Import and Export
##   'SPSS', 'Stata' and 'SAS' Files. R package version 2.1.1.
##   https://CRAN.R-project.org/package=haven
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files},
##     author = {Hadley Wickham and Evan Miller},
##     year = {2019},
##     note = {R package version 2.1.1},
##     url = {https://CRAN.R-project.org/package=haven},
##   }
citation(package = "psych")
##
## To cite the psych package in publications use:
##
##   Revelle, W. (2018) psych: Procedures for Personality and
##   Psychological Research, Northwestern University, Evanston,
##   Illinois, USA, https://CRAN.R-project.org/package=psych Version
##   = 1.8.12.
##
## A BibTeX entry for LaTeX users is
```

```
##
##   @Manual{,
##     title = {psych: Procedures for Psychological, Psychometric, and Personality Research},
##     author = {William Revelle},
##     organization = { Northwestern University},
##     address = { Evanston, Illinois},
##     year = {2018},
##     note = {R package version 1.8.12},
##     url = {https://CRAN.R-project.org/package=psych},
##   }
citation(package = "Hmisc")
##
## To cite package 'Hmisc' in publications use:
##
##   Frank E Harrell Jr, with contributions from Charles Dupont and
##   many others. (2019). Hmisc: Harrell Miscellaneous. R package
##   version 4.2-0. https://CRAN.R-project.org/package=Hmisc
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {Hmisc: Harrell Miscellaneous},
##     author = {Frank E {Harrell Jr} and with contributions from Charles Dupont and many others.},
##     year = {2019},
##     note = {R package version 4.2-0},
##     url = {https://CRAN.R-project.org/package=Hmisc},
##   }
##
## ATTENTION: This citation information has been auto-generated from
## the package DESCRIPTION file and may need manual editing, see
## 'help("citation")'.
citation(package = "DAAG")
##
## To cite package 'DAAG' in publications use:
##
##   John H. Maindonald and W. John Braun (2019). DAAG: Data Analysis
##   and Graphics Data and Functions. R package version 1.22.1.
##   https://CRAN.R-project.org/package=DAAG
##
## A BibTeX entry for LaTeX users is
##
```

```
##   @Manual{,
##     title = {DAAG: Data Analysis and Graphics Data and Functions},
##     author = {John H. Maindonald and W. John Braun},
##     year = {2019},
##     note = {R package version 1.22.1},
##     url = {https://CRAN.R-project.org/package=DAAG},
##   }
##
## ATTENTION: This citation information has been auto-generated from
## the package DESCRIPTION file and may need manual editing, see
## 'help("citation")'.
citation(package = "broom")
##
## To cite package 'broom' in publications use:
##
##   David Robinson and Alex Hayes (2019). broom: Convert Statistical
##   Analysis Objects into Tidy Tibbles. R package version 0.5.2.
##   https://CRAN.R-project.org/package=broom
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {broom: Convert Statistical Analysis Objects into Tidy Tibbles},
##     author = {David Robinson and Alex Hayes},
##     year = {2019},
##     note = {R package version 0.5.2},
##     url = {https://CRAN.R-project.org/package=broom},
##   }
citation(package = "e1071")
##
## To cite package 'e1071' in publications use:
##
##   David Meyer, Evgenia Dimitriadou, Kurt Hornik, Andreas
##   Weingessel and Friedrich Leisch (2019). e1071: Misc Functions of
##   the Department of Statistics, Probability Theory Group
##   (Formerly: E1071), TU Wien. R package version 1.7-2.
##   https://CRAN.R-project.org/package=e1071
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
```

```
##      title = {e1071: Misc Functions of the Department of Statistics, Probability
## Theory Group (Formerly: E1071), TU Wien},
##      author = {David Meyer and Evgenia Dimitriadou and Kurt Hornik and Andreas Weingessel and Friedrich Leisch},
##      year = {2019},
##      note = {R package version 1.7-2},
##      url = {https://CRAN.R-project.org/package=e1071},
##    }
```