DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

"I hate homework!" Could Musical Instrument Learning Help Reduce Academic Task Aversiveness?

Citation
Lui, Justin. 2023. "I hate homework!" Could Musical Instrument Learning Help Reduce Academic Task Aversiveness?. Master's thesis, Harvard University Division of Continuing Education.

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37374895

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

“I hate homework!” Could Musical Instrument Learning Help Reduce Academic Task Aversiveness?

Justin Lui

A Thesis in the Field of Psychology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2023

Copyright 2023 Justin Lui

Abstract

Academic tasks are often associated with negative emotions and have been shown to be a source of conflict between parents and children. Although many studies have been done surrounding academic task aversiveness (ATA), there is a lack of investigation into the activities that could potentially alleviate ATA. This study examines the relationship between musical instrument learning (MIL) and the potential mechanisms through which MIL can influence ATA – self-efficacy, conscientiousness, and negative error response. In addition, although not part of the original set of hypotheses, the variable family learning attitude was discovered as a key mediating variable between MIL and ATA.

Frontispiece

Illustration by Aleksei Naumov, 2021, istockphoto.com, license purchased.

Dedication

This study is dedicated to all the teachers who pour out their hearts to love and nurture the students who have been entrusted to their care. Your passion and love are contagious!

Acknowledgments

To my thesis director, Professor Richard J. McNally – Thank you for providing the crucial guidance in the development of my thesis. Despite your busy schedule, you've always made yourself available to review my work and to provide resources that serve as important references. Thank you for giving me the freedom to explore the topic of academic task aversiveness and its many correlates. During the multiple Covid lockdowns in Shanghai, even life's most basic necessities came into question. Thank you for your kind understanding and support during those difficult periods.

To my research advisor, Dr. Adrienne Tierney – Thank you for making the first steps of my thesis possible. I was self-doubting and did not even think I would make it through the thesis proposal process. Thank you for your heart-felt encouragement when I did not think I had what it took to make it through.

I'd like to thank my wife, Hannah. Thank you for supporting me throughout this master's degree journey. Thank you for being in my place while I spent hours studying and working on my thesis – from getting things done around the house to parenting our son. You are the best! I love you!

Lastly, a big thanks to my mom. You've always been supportive of me in everything I endeavored.

Table of Contents

List of Tables

List of Figures

Chapter I.

Introduction

Schoolwork is not always fun. Academic tasks can often be complex, require mental exertion, and create feelings of apprehension (Onwuegbuzie & Collins, 2001). Miller and Cohen (2001) described learning as effortful engagement, while Kool and Botvinick (2013) suggested that utilizing cognitive control is intrinsically costly and feels aversive. Given the types of negative emotions that can surround schoolwork, it is not surprising that many students feel stressed when faced with schoolwork. One study (Galloway et al., 2013) revealed that 56 percent of high school students considered homework a primary source of stress, whereas another (Pope, 2010) showed that more than 70 percent of students felt stressed by schoolwork. For younger children, the reluctance to do homework can also be a major source of conflict between parents and children (Kralovec & Buell, 2000; Vatterott, 2018).

Despite being associated with a multitude of negative emotions and a key source of contention in families, there has been a scarcity of research on activities that can help reduce academic task aversiveness (ATA). The present study examined the association between ATA and musical instrument learning (MIL), an extracurricular activity that can facilitate many positive attributes (Guhn et al., 2020; Sala & Gobet, 2017). This study investigated whether MIL could affect ATA through three mediating variables – self-efficacy, negative error response, and conscientiousness.

The present study used an online survey to collect data from 985 participants and took the form of an observational study. After the data were collected, Bayesian network analysis was performed to examine potential causal relationships between the variables and to detect potential collider variables. Structural equation modeling (SEM) was then used to perform parallel mediation analysis, and finally, traditional regression methods were used to verify the results.

Through the mediation analysis, I found that the proportion of effect accounted for by the hypothesized mediating variables was small. I then searched through the data of other collected variables and found that perceived family learning attitude accounted for a substantial portion of the variance even after applying the Holm-Bonferroni Method to correct for multiple comparisons.

Other interesting aspects of this investigation include the finding of multiple factors within the ATA measure (using exploratory factor analysis), the strong correlations that perceived family learning has with various variables, and the applying of bootstrapping method on multiple regression to overcome normality issues and control for potential confounds.

This literature review has been done in terms of the theoretical basis for this study as well as the constructs under examination. The literature review provided theoretical support for the hypothesis and motivation in terms of the importance of ATA. I learned that there is little research concerning activities that could help alleviate academic task aversiveness.

Theoretical Basis

The broaden-and-build theory (Fredrickson, 2001) suggests that people are in an optimal state to explore and learn new things when they are experiencing positive emotions. On the contrary, when people are experiencing negative emotions, their ability to learn and their level of interest in facing novel challenges are diminished. Many studies have shown that schoolwork is associated with a multitude of negative emotions such as frustration, annoyance, and tension (Trautwein et al., 2009). In addition, the aversiveness of academic tasks leads to conflict between parents and children (Galloway et al., 2013; Kralovec & Buell, 2000; Pope, 2010; Vatterott, 2018), which worsens negative emotions. It is therefore beneficial to identify the correlates, and if possible, the causes of those negative emotions to lessen them and help students flourish.

In the field of learning, the theory of identical elements proposed by Thorndike and Woodworth in 1901 suggests that the transfer of learning from one domain to another is possible, and that transfer is enhanced or restricted by the number of elements in common between the original context and the applied context (Hajian, 2019). The more similar the learning tasks are, the more the training in the original context will influence learning in the applied context. This study explores whether the benefits of learning musical instruments could transfer to the learning that occurs in formal schooling, and specifically in performing academic tasks. Prior studies have shown that elements such as self-efficacy, error response, and conscientiousness play a role in both musical instrument learning (MIL) and school education. These constructs were examined to determine whether they mediate the relationship between MIL and ATA and facilitate the transfer of benefits from one domain to another.

Variables Under Consideration

This section reports relevant research on constructs examined in this study, namely, musical instrument learning (MIL), the predictor variable, academic task aversiveness (ATA), the outcome variable, and the potential mediating variables error response, self-efficacy, and conscientiousness.

Academic Task Aversiveness

The term academic task aversiveness was first used by Solomon and Rothblum (1984) in academic procrastination research. They defined it as how unpleasant one finds school-related tasks; it is a strong predictor of procrastination (Milgram et al., 1988; Solomon & Rothblum, 1984). In the broader sense, academic task aversiveness is the result of the negative emotions surrounding mental exertion. Certain theories have been proposed to explain the phenomena of those negative emotions – cognitive dissonance theory, opportunity cost theory, and neurophysiological processes. In the area of learning theories, Cognitive Dissonance Theory explicitly focuses on the discomfort of learning. Festinger (1962) proposed that when individuals hold two contradictory beliefs or ideas, it is experienced as psychological stress, and effort is then made to resolve the contradiction and reduce the discomfort (Guzzetti, 1993). Empirical work prompted by the cognitive dissonance theory has shown that the stress brought on by cognitive dissonance can be measured both physiologically using skin conductance measurements (Croyle & Cooper, 1983) and psychologically using self-report measures (Elliot & Devine, 1994).

In the research area of self-control and willpower, Kurzban et al. (2016) stated that the performance of mental tasks is often shadowed by a sense of effort and that feeling is aversive. The authors proposed using an opportunity cost model to explain the negative affects associated with certain mental tasks. The model holds that certain computational mechanisms can only be deployed for a limited number of simultaneous tasks at any given moment. The sensation of effort reflects the forgoing of other opportunities in the pursuit of a certain task. The authors raised the question of why certain cognitive functions, such as general vision, carry no phenomenology of effort and those functions can be sustained continuously without deterioration of performance. Contrarily, tasks such as mathematical computations, vigilance tasks (flanker tasks), and "task-switching" paradigms – tasks that are highly related to academic tasks – are associated with the sensation of effort and decrease in performance over time. The model suggests that the aversive sensations are the aggregate of opportunity costs, and their magnitudes are directly proportional to each other. The emotional experience then acts as input to the decision-making process, often diverting attention from the task at hand.

From a neurophysiological perspective, human brains consume glucose, and strenuous cognitive activities require more glucose than simple ones (Ampel et al., 2018). Cheval et al. (2018) proposed the automatic process (as opposed to a controlled process) called energetic cost minimization, where individuals are drawn to the most cost-effective and energy-saving behaviors. Their experiments showed that when individuals were asked to perform avoidance tasks, human brains consume more energy avoiding inactive images than active ones, i.e., it is easier to reject energy-consuming tasks. The authors posited that, from an evolutionary perspective, saving energy has been beneficial for human survival as it allowed for more efficient behaviors in searching for food and shelter, competing for mates, and avoiding predators. It follows that mental exertion is not the default mode of operation and therefore requires conscious cognitive control, and hence a sense of effort. Many studies have also found that this aversiveness towards cognitive effort has a strong correlation with procrastination.

In Steel's (2007) meta-analysis, task aversiveness was a strong and consistent predictor of procrastination. In one of the first studies on academic task aversiveness, Solomon and Rothblum (1984) found that fear of failure and task aversiveness (rated on a 5-point scale) were most strongly related to procrastination, accounting for 49.4% and 18% of the variance respectively. In a related study, Milgram et al. (1988) studied task aversiveness by asking participants to rate their degree of perceived dysphoric affect on 54 everyday tasks (6 being academic). Aversiveness was significantly related ($r = 0.58$, $p < 0.01$) to behavioral delay and explained 33% of the variance in procrastination. Task aversiveness also had the highest correlation with procrastination among the variables, with task aversiveness having a correlation coefficient of .58, higher than covert negativism ($r = .49$) and perceived incompetence ($r = .25$).

Milgram et al. (1995) also studied academic task procrastination and asked respondents to rate their tendency to delay on 17 academic tasks. Items included: doing an assignment that requires independent work, buying school supplies, writing a term paper, keeping up with readings, and doing homework that requires a lot of reading. They found that the main effect between task aversiveness and delay was significant, $F(2,386) = 113.96$, $p < .001$. There was noticeably less delay when students dealt with pleasant tasks compared to neutral and unpleasant ones. In another study on task aversiveness and procrastination (Blunt & Pychyl, 2000), the researchers broke down the variables into finer components. They studied the procrastination of tasks in four stages (inception, planning, action, and termination) and task aversiveness in three categories (boredom, frustration, and resentment). Although all negative emotions were correlated with procrastination, boredom ranked the highest across all four task stages. These findings suggest that task aversiveness plays a major role in students' procrastination tendencies. What then, are the correlates and main causes of academic task aversiveness?

Aside from the correlation between academic task aversiveness and procrastination, researchers also studied the relationship between personality traits and task aversiveness. Lay (1992) studied the differences between trait procrastinators (individuals with a predisposition to procrastinate) and non-procrastinators. He found that trait procrastinators viewed their tasks as more aversive and viewed themselves as less competent than non-procrastinators. Even when controlling for negative affect, trait pessimism, and time management behavior, procrastinators still viewed their tasks as more negative. Lay and Brokenshire (1997) also found that conscientiousness was negatively correlated to task aversiveness and that task aversiveness mediated the relation between conscientiousness and behavioral delay. Notably, one of the six facets of conscientiousness is competence and, when rated by self, is similar to the measure of self-efficacy, which is also a hypothesized mediating variable in this study.

Aside from personality traits, several studies have also investigated the relationship between task characteristics and task aversiveness. Lonergan and Maher's (2000) findings reinforced Milgram et al.'s 1988 study and found that task autonomy (the opposite of imposition, in Milgram et al.'s study) was inversely correlated to task aversiveness. Senecal et al. (1997) found that participants had higher perceived task aversiveness when they knew that their performance would be evaluated. Furthermore, task aversiveness also increased when participants focused on performance rather than interest. Notably, performance focus is a construct that is related to another proposed mediating variable in this study, error response, discussed in a later section. Given the direct and indirect relationships that ATA has with conscientiousness, self-efficacy, and error response, could certain activities or training enhance these attributes and reduce ATA?

Musical Instrument Learning

Scientists have studied many different activities and training and their respective benefits for learners. One such area has received a substantial amount of attention – musical training (Sala & Gobet, 2017), and the studies have examined benefits that include cognitive skills, executive functions, personality traits, and academic performance.

In terms of cognitive abilities, Roden et al. (2014) studied students (7–8 years old) from different schools where some were offered musical instrument learning (MIL) and others were offered extra-curriculum natural science classes over a period of 18 months. They found that children who had musical training showed greater improvements in processing speed and auditory cognition. Separately, Janus et al. (2016) studied the effects of 20 days of musical summer camp training on cognitive abilities. Children (4–6 years old) who received musical training were able to achieve greater improvements in verbal fluency, sentence judgment, and visual search. In the area of executive functions, Jaschke et al. (2018) conducted an interventional and longitudinal study on the effects of musical training. The study was carried out over a period of 2.5 years among children with a mean age of 6.4 years old. The music groups performed better in inhibition on a Go/No-Go task, planning, and verbal intelligence compared to the control group. In a meta-analysis of 38 experimental studies on musical training, Sala and Gobet (2017) found that musical education had substantial effects on specific cognitive skills. They found that the average effect sizes for memory and intelligence were .34 and .35 respectively. However, not all areas benefited from musical training – some areas' effect sizes such as literacy were negative (i.e., the intervention group did not do as well as the control group). And for mathematics, phonological processing, and spatial processing, effect sizes were below .20, which were negligible. These intervention studies provide strong evidence that MIL could enhance performance in certain subject areas. These areas have been considered when designing the questionnaire for this study, for example, questions for error response and task aversiveness include math, science, writing, and memorization.

Although musical education in general has been shown to be beneficial, training involving musical instruments (vs. vocal, rhythm, or theory) yielded particularly strong results. In a population-level correlational analysis, Guhn et al. (2020) found that students who took musical instrument lessons did better academically. This study used educational records from 4 public schools in British Columbia, Canada, to examine the relationship between school music participation and academic achievement. The researchers found that the correlations with academic performance were stronger for instrumental music learning than vocal music. The difference in performance for English, math, and science resulted in effect sizes (Cohen's *d*) that ranged from .12 to .30, indicating small but significant differences. In addition, the study also found that highly engaged music students with multiple-year participation were academically more than one-year ahead of their peers. Guhn et al. (2020) obtained their results after controlling for an extensive set of potential confounds, e.g., prior academic achievement, social, economic, and cultural background variables, etc. Therefore, by controlling these potential confounds, even though the study was correlational, it provided compelling evidence on the impact of MIL. I have referenced their list of potential confounds when designing the questionnaire and considered the duration of learning when operationalizing the variable MIL.

Aside from cognitive functions and academic performance, researchers have also studied the relations between musical education and self-concepts. Costa-Giomi (2004) performed an intervention study where fourth-grade students were given a piano at home and weekly private lessons for three years. The results indicated that the piano students improved more in self-esteem and music class grades than did the control group. Although self-esteem and self-efficacy are different constructs, they are often correlated for students (Lane et al., 2002). This connection provides support for investigating the relationship between MIL and self-efficacy in this study. Another study on self-efficacy was done by Bugos and Cooper (2019), where a group of older adults received intense xylophone training for eight weeks. With results almost reaching standard significance, $F(1, 18) = 3.76$, $p = 0.068$, the musical training group had more improvement in general self-efficacy compared to the control group. The above review is by no means exhaustive and only provides a small sample of the studies that have been conducted on musical training. Most of the studies shine favorable light in terms of the effect of MIL, however, not all researchers agree that the benefits are due to the training.

Although much evidence favors the benefits of MIL, some critics claim that children who learn musical instruments inherently have higher intelligence and stronger executive functions. They argue that the higher intelligence and stronger executive functions afford the children more enjoyment when starting to learn an instrument and provide them with more advantages in continuing on their music learning journey (Sala

& Gobet, 2017). They further suggest that the observed superior academic performance is due to those inherent strengths and not MIL. However, Corrigal et al. (2013) performed a correlational study on MIL, academic performance, intelligence, and social-economic status. Although MIL was associated with parents' education and intelligence, the study found that MIL had a significant effect on academic performance even when intelligence was held constant, refuting the claim that differences were due solely to inherent characteristics.

In summary, researchers have conducted both interventional and correlational studies on MIL. The intervention studies demonstrated that MIL improves processing speed, language skills, executive functions, self-esteem, and memory. The correlation studies suggested that MIL is associated with stronger academic performance even when controlling for potential confounds such as intelligence and demographics. In the next sections, three attributes that could potentially connect MIL and ATA will be presented – error response, self-efficacy, and conscientiousness.

Error Response

Performing learning tasks such as solving math problems and memorizing information involves making errors and realizing errors have been shown to be aversive (Hajcak & Foti, 2008). Prior studies have shown that the realization of errors is accompanied by negative emotions such as fear, guilt, sadness (Zhao, 2011), anger, anxiety (Keith et al., 2020), and frustration (Keith & Frese, 2008). Since errors are an integral part of learning, how people react to errors (error response) could play a role in how aversive they view learning to be. For example, if students commit errors while doing math homework and they react negatively (anger, anxiety, or frustration) to those errors, they could come to associate such negative emotions with math homework, thus making it aversive. In the context of MIL, it is evident that MIL involves learning new musical pieces on a regular basis, which, inevitably is accompanied by making mistakes. People rarely achieve perfection on their first attempt. Is it plausible that, through repeated practice, MIL participants learn to handle errors better, and result in less ATA? There are two possible mechanisms: First, MIL could desensitize error-related negative affects; secondly, MIL could help develop a growth mindset that enhances error response.

In their study on dread, Berns et al. (2006) suggested the mechanism of anticipatory adaptation — by rehearsing a painful event and exciting the pain network repeatedly, the brain could be desensitized to the noxious stimuli of the actual event. Repeatedly encountering errors while practicing a musical instrument could provide the rehearsal of handling errors. In another area of research, it is also well established that cognitive behavioral therapy (CBT) uses desensitization to effectively treat negative emotional cues such as traumas and phobias (de Roos et al., 2011; Leutgeb et al., 2009). Repeatedly making mistakes in the process of MIL could therefore be a form of desensitization for errors, which in turn reduce negative arousal. In addition, a meta-analysis has shown that error management training transfers well from one context to another (Keith & Frese, 2008). Could the desensitization of error-induced emotions in MIL transfer to a general academic setting? This is a question that the present study aimed to answer.

The second potential mechanism through which MIL could enhance error response is mindset. Prior research has shown that MIL enhances the growth mindset (Müllensiefen & Harrison, 2020), while growth mindset enhances error management in

MIL (Kruse-Weber & Parncutt, 2014). There is also research that links growth mindset to event related potentials when committing errors, which are brain signals immediately following an event that can be measured by electroencephalogram (EEG). One study (Moser et al., 2011) showed that individuals with a stronger growth mindset had stronger error positivity signals (Pe) and better attention focus after committing an error. Contrarily, fixed mindset individuals had stronger P3 signals (Mangels et al., 2006), which are related to surprise and outcome. This finding was coherent with Senecal et al.'s (1997), where participants focused on performance showed higher task aversiveness. And in line with prior mindset research, growth mindset individuals are more focused on mastery and learning the correct response, whereas fixed mindset individuals are more focused on performance and outcome. Since MIL is related to mindset and mindset is related to error response, it is plausible that MIL could affect error response. Therefore, error response was investigated as a mediating variable in this study.

Self-Efficacy

The second variable that could link MIL and ATA is self-efficacy. As presented above, prior studies have shown a significant effect of MIL on self-esteem (Costa-Giomi, 2004), which is related to self-efficacy among students (Lane et al., 2004), thereby providing support for the link between MIL and self-efficacy. Also, in a study conducted with twenty older adults with results almost reaching significance, $F(1, 18) = 3.76$, $p = 0.068$, musical instrument training potentially increased general self-efficacy (Bugos & Cooper, 2019). In terms of the correlation between self-efficacy and task aversiveness, the studies performed by Milgram et al. (1995) and Lay (1992) showed significant correlations between perceived task capability (a measure that is related to self-efficacy) and task aversiveness. Students considered themselves more capable of doing academic tasks that were more pleasant. These studies suggest that self-efficacy could potentially be a factor that affects how MIL relates to task aversiveness.

Conscientiousness

Another potential mechanism through which MIL can act on ATA is personality traits. Conscientiousness is a dimension of the Big Five personality traits (Costa & McCrae, 1992). Persons high in conscientiousness are described by phrases such as "a reliable worker," "perseveres until the task is finished," and less so by phrases such as "lazy", and "easily distracted" (John et al., 1991). Learning a musical instrument requires regular practice, which in turn requires self-discipline and perseverance (Covay & Carbonaro, 2010). Engagement in MIL could therefore enhance conscientiousness. Although some studies have shown personality traits to be rather stable (Roberts et al., 2006), other studies have shown that they are malleable, especially during childhood (Soto & Tackett, 2015). In fact, the study by Hille and Schupp (2015) showed that students who undertook musical training have conscientiousness measures that were .23 ($p < .05$) standard deviations higher than their non-trained counterparts. Although the study used observational data, this difference was obtained after propensity score matching for an extensive list of potential confounds such as social economic status (SES), parents' education, personality, migration background, musical inclinations, amount of contact with school, disposition to help with homework, whether child shares a room with siblings, etc. The study offered strong support that undertaking MIL can influence the personality traits of children and adolescents, especially in the dimension of conscientiousness.

Although there is no research testing the direct relationship between conscientiousness and ATA, conscientiousness has been shown to be negatively correlated to procrastination (Lay & Brokenshire, 1997) and positively correlated academic achievement (Altanopoulou & Tselios, 2018; Schniederjans & Kim, 2005), both of which are strong correlates of ATA (Goroshit & Hen, 2021; Milgram et al., 1995). In addition, one study (Lay & Brokenshire, 1997) has shown association between conscientiousness and perceived task pleasantness for the area of job search tasks. Although their similarity with academic tasks is unknown, the study suggests that conscientiousness and ATA could be associated. Therefore, conscientiousness was also investigated as a potential mediator in this study.

A review of the literature on MIL and ATA indicates that there is currently a lack of study on the relationship between the two. At the same time, prior research supports the notion that there could be association between MIL and ATA through various mediators. To evaluate this relationship, the present study examined the mediating roles of error response, self-efficacy, and conscientiousness between MIL and ATA.

Chapter II.

Research Methods

This study took the form of an observational study and used an online survey to collect data for analysis. The recruiting method and survey were approved by the Harvard University Area Institutional Review Board. The following sections describe the source of participants and the measures used.

Participants

Since one of the key parameters invested in this study was academic task aversiveness (ATA), full-time students were the target participants. The number of participants required for the study was estimated using G*Power based on the type of analysis and effect size. Some effect sizes of MIL from prior studies are as follows. In Jaschke et al.'s (2018) study of MIL and executive functions, the paper provided F values only and not effect size. Using Thalheimer and Cook's (2009) formulas and Jaschke et al.'s information (F = 4.46, number of intervention subjects (Nt) = 80, number of control subjects (N_c) = 66), the calculated Cohen d was .353 for MIL on inhibition. Costa-Giomi (2004) studied self-esteem using weekly private piano lessons for three years. Again, only F values were provided. Using the formulas provided by Thalheimer and Cook (2009) and the values from Costa-Giomi (2004) (F = 11.16, N_t = 67, N_c = 50), Cohen d was .384. In a correlational study (Guhn et al., 2020), effect sizes ranged from .36 to .46. Conservatively, I chose the number .38 for power analysis. Using multiple linear regression as the analysis method and .38 as the correlation coefficient, $f^2 = .169$, G*Power provided a total required sample size of 95 ($f^2 = R^2 / (1-R^2)$).

Since the proportion of students who continue to learn musical instruments reduces drastically upon entering college, high school students were the preferred group for recruitment. Originally, the plan was to recruit high school students from international schools in Shanghai, where I reside. However, due to COVID-19 lockdown at the time of recruitment, the response rate was extremely low (10%), and the total number of returned surveys could not reach the required power for analysis. Therefore, multiple online crowdsourcing websites were reviewed and compared to recruit participants instead. Based on the recommendations for sourcing quality participants on apa.org (Palmer & Strickland, 2016) several measures were taken in the recruiting and survey design process. These measures included keeping the survey duration to a minimum and providing an accurate estimate of the job description, attention, English proficiency and bot checks, and compensating participants in line with market rate. After conducting two pilot studies ($N = 100$), one on mturk.com and the other on prolific.co (Prolific), it was found that Prolific had a lower percentage of participants who got both attention check questions incorrectly. Thus, Prolific was used to recruit participants. Participants were limited to the US, Canada, and the UK. The number of participants recruited was 985, of which 910 completed the study. Out of the 910 participants who completed the study, 10 were duplicates, 16 failed both attention checks and 9 did not pass the English and bot test and were excluded from the study. The average duration to complete the survey was 16.8 minutes with a standard deviation (*SD*) of 6.7 minutes. No participants were excluded because their duration was too short (under 3 *SD*). The final number included in the analysis was 875, with 530 females and 345 males. The average age among participants was 19.3 years old with a standard deviation of .99. Grade distribution was as follows: grade 11, 0.2%; grade 12, 9.0%; college freshman, 26.4%; college sophomore, 42.2%; college junior, 11.5%; college senior, 3.5%; part-time student 7.1%. Ethnicity distribution was as follows: African, 9.4%; Asian, 17.8%; Caucasian. 59.7%; Hispanic, 7.9%; other ethnicities, 5.2%.

Instruments

Participants recruited on Prolific were directed to Harvard's Qualtrics website to complete the questionnaire. The questionnaire comprised 157 questions in total with 21 conditional items (i.e., they only appeared if participants answered yes or no to certain questions). The average time to complete the survey was 16.8 minutes. There were 9 parts to the survey and the ones after demographics were randomized by section to counter survey bias. The 9 parts of the survey were: informed consent, demographics, MIL, error response, growth mindset, ATA, self-efficacy, conscientiousness, and sports and performing arts.

Musical Instrument Learning

The MIL section of the questionnaire collects information on whether the participant has (or has had) musical education and various aspects of it. Referencing the population study by Guhn et al. (2020), the MIL questionnaire included information on the duration, type of musical education, participation in school band or orchestra, attendance of private lessons, and the highest level music examination taken. Details of the questionnaire can be found in Appendix 1. In this study, the variable MIL was operationalized in two ways. The first categorized MIL participants as those who have had 8 or more years of musical instrument training, either private lessons or school band/orchestra, and that they are either currently engaged in MIL or have only stopped within the last year. The second method differs from the first in categorizing MIL participants as those with 8 to 11 years of training. The rationale for the second operationalization method will be presented in the analysis section. The minimum duration of MIL, 8 years, references the study performed by Guhn et al. (2020), which was a population-level study that showed significant differences for MIL participants across multiple academic subjects.

Error Response

Five academic scenarios were created for this study. Each scenario involves committing an error in an academic task with varying degrees of consequences in terms of the amount of time required to correct the mistake (10 minutes to 2 hours). To check for successful manipulation, I tested the average differences between error scenarios using *t*-tests. The arousal of negative emotions when realizing a smaller and less time-consuming mistake should be lower than realizing mistakes that take longer time to correct. As such, it is hypothesized that Scenario 3 (“You worked really hard over the weekend to complete an assignment that you thought was due on Monday. Then you came to class, and now you realized that the assignment is actually due a week later.”) will trigger a lower negative error response than Scenario 1 (“You just spent 10 minutes working on a math problem. Now, you realized you made a mistake in the first few steps and have to redo the whole solution again.”), which in turn will trigger a lower negative error response than Scenario 5 (“You are writing a 5-page essay for English literature/language arts class. You are almost finished and are writing the ending section. Now you realize a couple of the previous sections aren't coherent and have to rewrite some portions of the review. This is going to take you another 2 hours.”). Welch two sample *t*-test between Scenario 3 and Scenario 1 showed that Scenario 1 does indeed trigger a higher average negative emotional response, $t(1574.5) = 26.03$, $p < .001$. In addition, *t*-test between Scenario 1 and Scenario 5 shows that Scenario 5 also triggered a higher average negative emotional response than Scenario 1, $t(1797.4) = 6.11$, $p < .001$. Hence, it was concluded that manipulation of the error scenarios was successful.

Following other studies on error response (Zhao, 2011; Keith et al., 2020), various Positive and Negative Affect Scales (PANAS) were compared to decide which was the best to use in this study. The International PANAS Short Form (I-PANAS-SF) (Thompson, 2007) was chosen because it was developed with non-native English speakers in mind, which was well suited for the original population under study at the international school in Shanghai. In addition, the I-PANAS-SF is also more concise (10 items vs. 20 items in the regular PANAS) and can potentially reduce respondent fatigue since the I-PANAS-SF needs to be answered for each of the 5 error scenarios. When evaluated by Thompson, the scale had Cronbach alphas of .78 and .76 for the positive and negative affect subscales respectively. For this study, the intra-scenario Cronbach alphas for the 5 scenarios ranged from .704 to .779 indicating that internal consistency for the I-PANAS-SF was acceptable (Tavakol & Dennick, 2011). The error response scenarios are included in Appendix 2 and the I-PANAS-SF scale is included in Appendix 3.

Self-Efficacy

The instrument used to measure self-efficacy was the General Measure of Self-Efficacy (GMSE) developed by Sherer et al. (1982). The scale was developed with a total of 36 items. After performing factor analysis, the results yielded a two-factor solution, one being general self-efficacy and the other being social self-efficacy. The general self-efficacy scale was used in this study and had a Cronbach's alpha reliability coefficient of .86. The construct validity of the self-efficacy scale was also verified by Sherer et al. by comparing the scale with other established relevant scales, such as the Internal-External Control Scale, Ego Strength Scale, and Self-Esteem Scale (Sherer et al., 1982). For this study, the Cronbach alpha was .899, indicating an acceptable internal consistency (Tavakol & Dennick, 2011). The GMSE is included in Appendix 4.

Conscientiousness

The measure used for conscientiousness was the conscientiousness scale in the Big Five Inventory (John et al., 1991; John et al., 2008; Benet-Martínez & John, 1998) (Appendix 5). The conscientiousness scale has 9 items, 4 of them being reverse-scored items. This scale was chosen because it is relatively short and available for non-commercial research purposes (*The Big Five Inventory*, 2009). According to Benet-Martínez & John (1998), the scale correlates well with the NEO PI-R (Costa & McCrae, 1992)mean r = .75) and has a mean Cronbach alpha of .83. The Cronbach alpha for the conscientiousness scale in this study was .84, indicating internal consistency was maintained and acceptable (Tavakol & Dennick, 2011). The survey includes 6 dummy items from other parts of the Big Five Inventory to prevent participants from guessing what the scale was trying to measure.

Academic Task Aversiveness

The topic of ATA has been studied by multiple researchers (Milgram et al., 1995; Milgram et al., 1988; Solomon & Rothblum, 1984). However, there is no widely accepted scale for measuring ATA. The ATA scale used in this study consisted of academic tasks used by Solomon and Rothblum (1984) plus those concerning different subjects, namely, language arts, math, and science. Tasks concerning different subjects have been added because studies (Sala & Gobet, 2017) have shown that MIL has varying degrees of correlation with different subjects. Therefore, collecting this data will facilitate analysis of MIL with the aversiveness of different subject tasks. Details of the questionnaire are included in Appendix 6. The Cronbach alpha for the 8-item ATA is .456, indicating a lack of internal consistency and the possibility that it could be measuring more than one construct. This will be addressed by exploratory factor analysis in the analysis section.

Chapter III.

Results

This section consists of a brief introduction of Bayesian network analysis followed by the analyses and results of this study. Analyses for this study are composed of 3 main parts: Bayesian network analysis, structural equation modeling (SEM) for parallel mediation analysis, and traditional regression methods. Foster (2010) suggested that the best practice for causal inference involves using multiple approaches to estimate the effect. And if the effects are consistent across different analytical approaches, then one can be more confident in the findings. That is the reason why Bayesian network, SEM, and regression methods have been used in this study.

Bayesian Networks Introduction

Traditionally, psychology researchers have shied away from making causal suggestions based on observational data, and for good reason. It is dangerous to make causal claims based purely on correlational data (Foster, 2010). However, no intervention or treatment can be implemented without causal reasoning, and errors in analysis can occur without clarity in causal direction (e.g., mistaking a collider or mediating variable as a confounder variable). With advances in causal theories (Pearl, 2016) and analytical tools (McNally, 2016), it is no longer productive to completely ignore causal inference in observational studies. One technique in causal inferencing is Bayesian networks.

Bayesian networks are directed acyclic graphs (DAGs) that contain nodes and arrows. The nodes represent variables and the arrows represent probabilistic dependencies between the variables (Scutari, 2010). Theoretically, DAG is a factorization of the joint probability distribution of the node set into a set of local probability distributions, where factorization is based on the Markov property. The Markov property states that if two nodes are not connected, they are not directly dependent on each other (Scutari, 2010). The DAG provides an estimation of causal dependencies among the variables (McNally et al., 2022) and "aspirationally" describes the direction of causation (McNally, 2016). The arrows in a DAG represent the direction of prediction or directional dependence relations. The presence of a descendant more strongly implies the presence of the parent, rather than vice versa. In other words, the lines show conditional dependence associations or probabilistic dependencies among the variables.

One benefit of using Bayesian network analysis is that it allows for simultaneous investigation of a high number of relational models (Silas et al., 2022). Traditional hypothesis testing uses a priori knowledge and limits the number of possible models tested. Since Bayesian network does not rely on a priori knowledge, it shows dependency relationships that one might not consider. For example, it could indicate potential mediators and colliders, thereby preventing conditioning on a mediator or collider mistakenly, which can eliminate valid relationships or create spurious relationships (Foster, 2010).

Bayesian networks have been used in multiple areas of psychology research including developmental psychology (Foster, 2010) and clinical psychology (McNally et al., 2022) where intervention studies are often not possible (Silas et al., 2022). In most cases, causal claims are not possible with Bayesian network modeling because assumptions are not verifiable, namely causal faithfulness and causal sufficiency Field (Briganti et al., 2022). Two additional assumptions are required when building a Bayesian network: Firstly, no bidirectional relationship and loop structures are present in the network; and second, no missing important variables. It is partly due to the second assumption that many variables, other than the core variables, have been included in this study (e.g., demographics, perceived autonomy, opportunity cost, family learning attitude, etc.).

In this study, the software used for performing Bayesian network analysis is the bnlearn package in R (Scutari, 2010). bnlearn uses a two-step approach to create a Bayesian network. The first step uses a learning algorithm to determine the structure of the network, i.e., the placement of the arrows. The second step estimates the parameters of the local distribution functions (Scutari, 2010). Structure learning is conceptually based on two frameworks: d-separation (directional separation) and v-structures (Scutari, 2010). D-separation is used to remove edges (Silas et al., 2022): When X and Y are correlated but become uncorrelated when conditional on Z, it implies that X and Y do not have direct effect on each other, and the connection removed. V-structures are used to determine the direction of prediction: When X and Z are not significantly correlated, but X and Y are correlated, and Z and Y are correlated, and X and Z become correlated when conditional on Y, it implies that X “causes” Y and Z “causes” Y; thus arrows point from X to Y and from Z to Y (Silas et al., 2022). The results of Bayesian network analysis will be presented in a later section. Presented next is the treatment for the variable ATA.

Exploratory Factor Analysis for ATA

As mentioned in the instrument section, the ATA scale is composed of items that were used in prior studies (Milgram et al., 1995) as well as tasks that were added to include different subject areas, namely, writing, history, math, and physics. The Cronbach alpha for the 8-item ATA measure was .456, indicating a lack of internal consistency and that the various items likely measured more than a single construct. Therefore, exploratory factor analysis was used to determine whether the scale should be decomposed into several subscales. Using fa.parallel in the R Psych package (Revelle, 2022), a parallel analysis was performed and resulted in the scree plot in Figure 1.

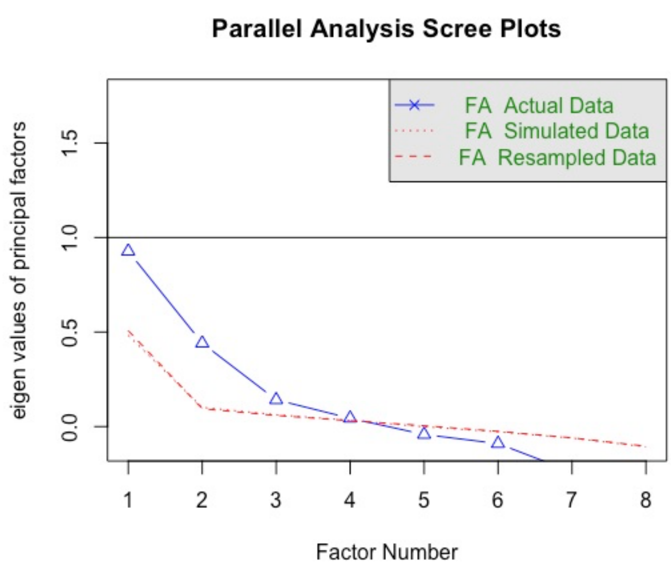

Figure 1. Parallel Analysis Scree Plots

Parallel analysis used to determine starting number of factors for exploratory factor analysis.

Observing that the eigenvalues (blue) start to plateau at 3 factors and the point of inflection at 4 factors, choices of 3 or 4 factors would be good starting points for the exploratory factor analysis (Dang, 2021). Assuming there is a certain degree of correlation between the factors, oblique rotation (oblimin) was used. Again, using the fa function in the psych package in R and a 3-factor model, the resulting root means square of residuals (RMSR) was .01; the root mean square error of approximation was .014; and the Tucker-Lewis Index was .987 indicating that the 3-factor model provided an acceptable fit (MacCallum et al., 1996; Schermelleh-Engel et al., 2003). The loadings for the items are shown in Table 1. None of the items loaded onto more than 1 factor with a loading greater than 0.3. The 3 factors have been named according to the cognitive nature of the tasks: factor 1, high cognitive demand; factor 2, low cognitive demand; and factor 3, memorization. The high cognitive demand scale was chosen for the rest of the analysis due to the higher number of survey items (3 vs. 2) and higher Cronbach alpha (.50 vs. .29). The Cronbach alpha remained lower than the recommended value of .70 (Tavakol & Dennick, 2011); however, that could have been due to the low number of items in the subscale. The zero-order correlations table for all the variables is presented in Table 2 below.

Table 1. Exploratory Factor Analysis for ATA.

Loadings	Factor 1 – high cognitive demand	Factor 2 – low cognitive demand	Factor 3 – memorization
Writing a term paper		.680	
Studying for an exam	.417		
Keeping up with weekly reading assignments			
Performing administrative tasks		.324	
Attending meetings for team projects			
Practicing math problems	.674		
Memorizing historical facts			.466
Learning new physics concepts	.426		

Table 2. Zero-Order Correlations Table.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Age	1																
2. Family learning attitude	-0.04	1															
3. Academic performance	-0.09	0.08	1														
4. MIL	-0.05	0.25	0.12	1													
5. Error Response Neg	0.12	-0.03	-0.06	-0.06	1												
6. Error Response Pos	0.01	0.2	0.07	-0.05	0.2	1											
7. Growth mindset	0.04	0.05	0.06	0	-0.07	0.03	1										
8. ATA	-0.02	-0.22	-0.04	-0.12	0.15	-0.3	-0.01	1									
9. Perceived opportunity cost	0	-0.22	-0.02	-0.11	0.23	-0.19	-0.03	0.39	1								
10. Perceived autonomy	0.01	0.19	0.07	-0.01	-0.06	0.22	0.07	-0.32	-0.24	1							
11. Self-efficacy	-0.03	0.28	0.19	0.03	-0.28	0.25	0.12	-0.15	-0.3	0.26	1						
12. Conscientiousness	0.05	0.23	0.27	0.04	-0.12	0.22	0.07	-0.06	-0.19	0.18	0.66	1					
13. Sports and performing arts participation	0	0.06	0.02	0.15	0.01	0	-0.03	0.01	-0.03	-0.07	0.12	0.06	1				
14. Fraternal ethnicity	0.07	-0.04	0.02	-0.07	0.04	0.05	-0.01	0	0.03	0.03	0.06	0.03	0.02	1			
15. Fraternal education	-0.06	0.32	0.06	0.21	-0.07	-0.01	0.01	-0.06	-0.07	0.02	0.03	0.06	0.09	-0.05	1		
16. Maternal ethnicity	0.07	0.01	-0.03	-0.11	0.04	0.03	0	-0.02	-0.03	0.05	0.06	0.04	-0.03	0.64	-0.08	1	
17. Maternal education	-0.04	0.26	0.07	0.23	0.04	0	-0.01	-0.01	-0.01	-0.04	-0.02	0.04	0.1	-0.03	0.59	-0.05	1

Bayesian Network Analysis 1

The package bnlearn in R (Scutari, 2010) version 4.7.1 was used to implement Bayesian network analysis for this study. The sample code from Briganti et al. (2022) was adopted with minor modifications to correct labeling errors in the original code. The final code used for this study is included in Appendix 7. Parameters for the analysis were set following the guideline in Briganti et al. (2022) and are included below. The number of bootstrap iterations was set at 1000. Certain relationships were whitelisted and blacklisted according to prior research. The lists along with references are included in Table 3. Arcs that were whitelisted in one direction were suggested by prior interventional studies, whereas arcs that were whitelisted in both directions were suggested by observational studies. Arcs from all measured variables to inherent variables were blacklisted (e.g., self-efficacy to maternal education).

Table 3. Whitelist and Blacklist Table

Blacklisted Arcs

From	To
Age, Family Learning Attitude, Academic Performance, MIL, Negative Error Response, Positive Error Response, Growth Mindset, ATA, Perceived Opportunity Cost, Perceived Autonomy, Self-Efficacy, Conscientiousness, Sports and Performing Arts.	Fraternal Ethnicity, Fraternal Education, Maternal Ethnicity, Maternal Education

Uni-Direction Whitelisted Arcs

From	To	References
MIL	Growth Mindset	Mullensiefen & Harrison, 2020; Holochwost et al., 2021
MIL	Self-Efficacy	Bugos et al., 2016; Costa-Giomi, 2004

Bi-Direction Whitelisted Arcs

Between		References
MIL	Academic Performance	Hille et al., 2015; Guhn et al., 2020
MIL	Negative Error Response	Keith & Frese, 2008; Kruse-Weber et al., 2015
MIL	Conscientiousness	Hille & Schupp, 2014
ATA	Academic Performance	Goroshit & Hen, 2021
ATA	Self-Efficacy	Milgram et al., 1995; Lay, 1992
ATA	Perceived Autonomy	Lonergan & Maher, 2000
ATA	Perceived Opportunity Cost	Kurzban et al., 2013
Self-Efficacy	Sports/Performance Arts	Milgram et al., 1995; Lay, 1992
Growth Mindset	Self-Efficacy	Bai et al., 2021
Negative Error Response	Self-Efficacy	Trice et al., 1991
Growth Mindset	Negative Error Response	Kruse-Weber et al., 2015

Arcs in Bayesian analysis are whitelisted and blacklisted according to associated references.

The Grow-Shrink algorithm (Margaritis, 2003), a constraint-based algorithm that finds conditional independence relationships using statistical tests, was found to be the best learning algorithm that provided the most relevant relationships for our variables of interest, namely MIL and ATA. Arc strength threshold was set at .85 and direction criteria was set at .50. Edge strengths were evaluated using Bayesian Information Criteria (bic-g). The directed acyclic graph (DAG) produced using these parameters is shown in Figure 2. The direction of the arrows indicates potential causality, while the thickness and darkness of color indicate the strength of the association. MIL points to growth mindset, conscientiousness, academic performance, and weakly at self-efficacy and negative error response, indicating that learning a musical instrument influences one's growth mindset, conscientiousness, academic performance, self-efficacy, and negative error response. Academic performance weakly affects academic task aversiveness. Although the DAG shows MIL's influence on the mediating variables, the hypothesized relationship with ATA was not confirmed by this analysis.

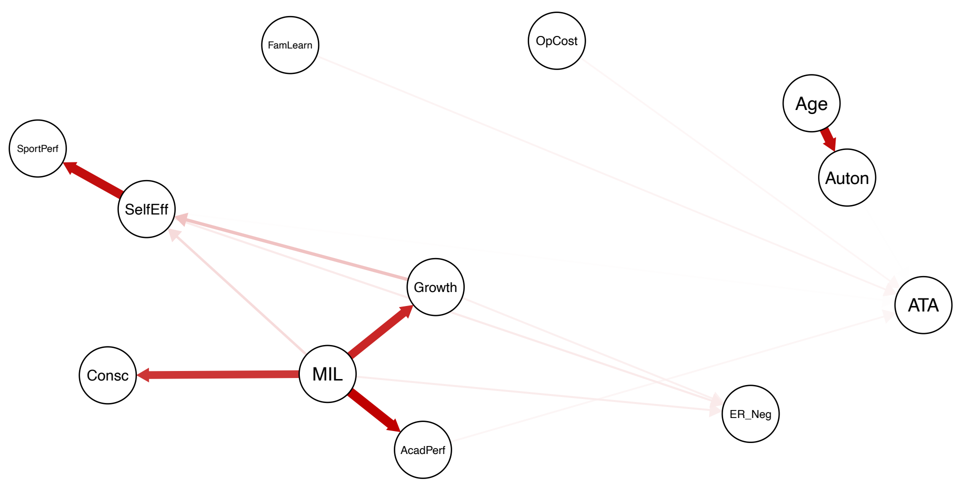

Figure 2. Directed Acyclic Graph 1.

DAG produced using Grow-Shrink algorithm in bnlearn in R.

Mediation Analysis Using Structural Equation Modeling 1

The next part of the analysis used structural equation modeling to investigate the hypothesized mediation relationships. The package Lavaan (Rosseel, 2012) was chosen because it allowed for parallel paths as well as a discrete predictor. Lavaan finds the best fit for a model by minimizing the model covariance and the sample covariance. The path diagram is presented in Figure 3.

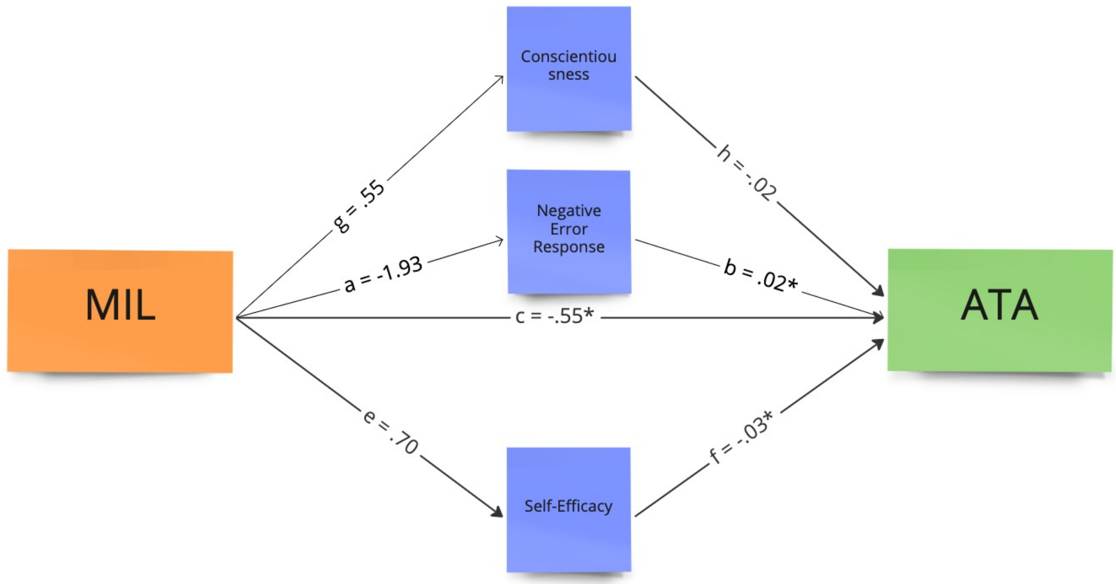

Figure 3. SEM Path Diagram 1.

Path coefficients using structural equation modeling with the Lavaan package in R. Indirect effects b_{ab} = -.03, p = .25; b_{ef} = -.02, p = .56; b_{gh} = .01, p = .59. *Total effect* $b_c + b_{ab} + b_{ef} + b_{gh}$ = -.60, p = .01. ** denotes* $p < .05$.

SEM showed that MIL did not have significant effects on the mediating variables and that the indirect effects were also insignificant. Only MIL had a significant direct effect on ATA, b_c = -.55, $p < .05$. Since prior studies have shown correlations between MIL and the mediating variables, I searched through the literature to determine whether the way MIL was operationalized could affect the analysis outcome. It was found that the duration of MIL indeed could influence various abilities (Daly & Hall, 2018), thus, its relations with the mediating variables were investigated.

Effects of Musical Instrument Learning Duration

In most musical training studies, the variable has been treated as dichotomous (Corrigall et al., 2013; Costa-Giomi, 2004; Guhn et al., 2020); one group received musical training for a certain number of years compared to another group that did not receive training. However, researchers Daly and Hall (2018) demonstrated that the length of training correlated to various skill and performance levels. Using ggplot in R, the mediating variables were curve-fitted against MIL duration and are shown in Figure 4.

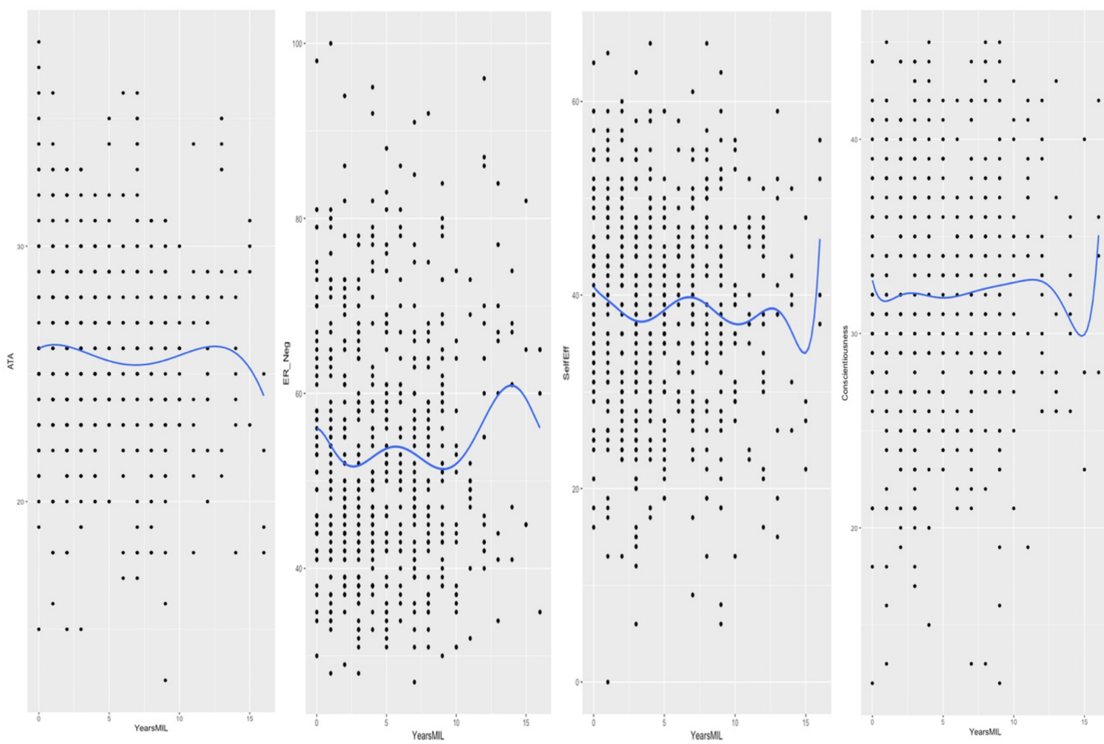

Figure 4. Mediating Variables vs. Years of MIL.

Curve fitting mediating variables against years of MIL.

The graphs show that all the dependent variables behave exceptionally at extreme training durations (13-15 years). At first sight, this was counterintuitive. If a training is good, shouldn't more training be better? However, after conducting some literature search, the reasons for the sudden change in effect direction in the graphs above became apparent. The common curriculums for MIL end after 10-12 years (*ABRSM*, n.d.; "Certificate of Merit®," n.d.). Near the end of the curriculums is training to prepare students for performances and competitions, similar to the level of a professional musician. Studies have found differences between amateur and professional musicians in terms of personality (Kuckelkorn et al., 2021), intrinsic versus extrinsic motivation (Juniu et al., 1996), and even brain structures (Gaser & Schlaug, 2003). Kuckelkorn et al.'s study (2021) found that professional musicians (compared with amateur musicians) were associated with higher neuroticism, lower agreeableness, and lower conscientiousness. Juniu et al. (1996) found that professional musicians had a higher association with extrinsic motivation and perception of work rather than leisure. In addition, another study found that in a work environment where performance is highly valued (e.g. a professional practice), negative error response was positively correlated with performance (Zhao, 2011). In a musical performing environment where mistakes are rarely tolerated, professional level musicians could thus be associated with higher negative error responses. It is possible that the aforementioned effects would also lessen the benefits that MIL has as students progress beyond the amateur years of training. With these findings, the MIL group criteria was adjusted to 8-11 years of learning. Unfortunately, this resulted in a significant reduction in the number of participants in the MIL group (N = 48), which impacted the power of the following analyses. Furthermore, the generalizability of the results was also reduced. These issues will be addressed in the future research section.

Bayesian Network Analysis 2

The package bnlearn in R (Scutari, 2010) was again used to perform Bayesian network analysis using data for the new MIL group that only included participants with 8-11 years of training. Whitelist, blacklist, and other parameters were kept constant as in previous analysis. The Grow-Shrink algorithm produced the DAG in Figure 5, where MIL points to growth mindset, conscientiousness, academic performance, and weakly at self-efficacy, indicating that learning a musical instrument influences one's growth mindset, conscientiousness, academic performance, and self-efficacy. In turn, growth mindset and self-efficacy affect ATA. This supports the hypothesis that MIL's effect on ATA is mediated by self-efficacy. From this network model, I learned that academic performance is a collider variable (Briganti et al., 2022), as it is influenced by MIL and weakly by ATA. If academic performance were held constant by mistake when performing a regression analysis between MIL and ATA, it could introduce spurious association between MIL and ATA (Silas et al., 2022). The network model also indicates that sports and performing arts is a confounding variable for conscientiousness and negative error response. If one were to analyze the relationship between those two variables, one should hold sports and performing arts constant to eliminate the confounding effect. Another interesting feature in the network model is v-structure, where a predictor variable and a third variable both point to the outcome variable, but the predictor variable and third variable are not associated. According to Pearl (2016), in a v-structure where the predictor variable and the third variable only become associated when conditioned on the outcome variable, that would indicate causal effect of the predictor on the outcome variable. In the DAG, opportunity cost (OpCost) and self-efficacy (SelfEff) both point to ATA. However, opportunity cost and self-efficacy have no connection between them. If they become correlated when holding ATA constant, it would suggest that self-efficacy have a causal effect on ATA. This v-structure relationship will be verified using partial correlations method in a later section.

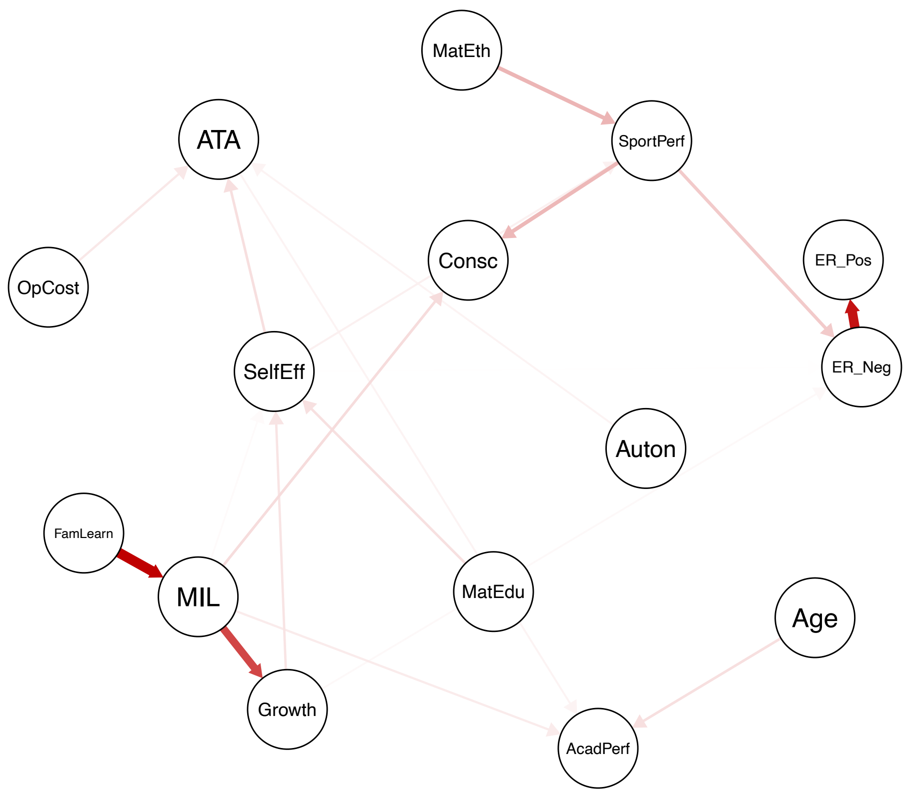

Figure 5. Directed Acyclic Graph 2.

DAG generated using the Grow-Shrink algorithm in bnlearn in R. MIL operationalized with 8 to 11 years of training.

Mediation Analysis Using Structural Equation Modeling 2

The package Lavaan (Rosseel, 2012) was used again for the mediation analysis. Figure 6 presents the path diagram and the coefficients for the model. The path coefficients reflect the connection strengths and represent the response of a dependent variable to a unit change in the independent variable when other variables are held constant (Bollen, 1989). Path coefficients and regression coefficients are similar – a positive coefficient implies that a unit increase in the independent variable will lead to an increase in the dependent variable proportional to the path coefficient.

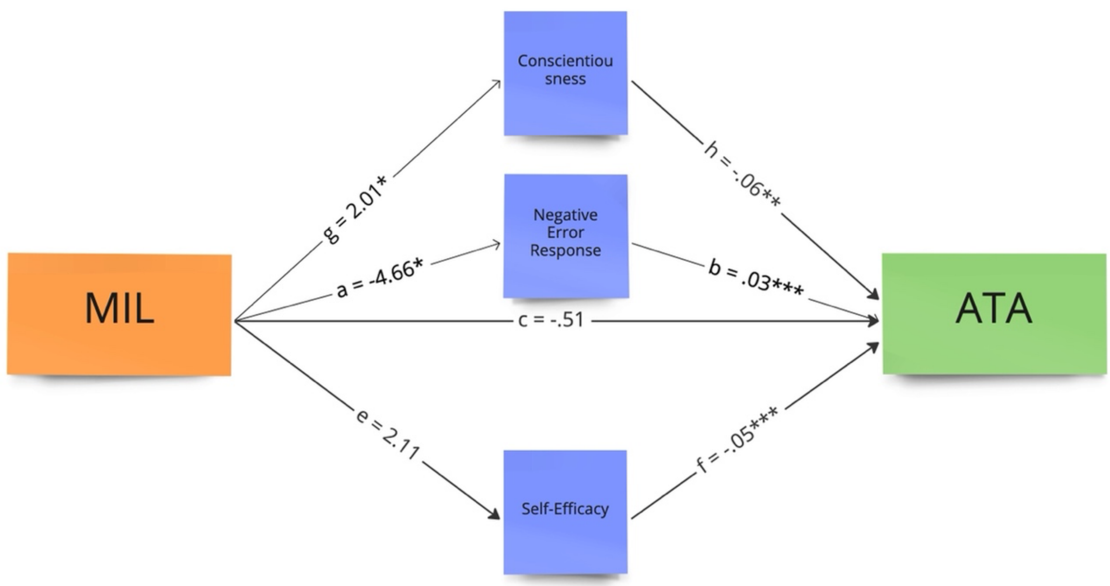

Figure 6. SEM Path Diagram 2.

Path coefficients using structural equation modeling with the Lavaan package in R. Indirect effects: $b_{ab} = -.14$, $p < .05$; $b_{ef} = -.11$, $p = .26$; $b_{gh} = -.11$, $p = .14$. *Total effect* $b_c + b_{ab} + b_{ef} + b_{gh} = -.87$, $p < .01$. ** denotes* $p < .05$*, ** denotes* $p < .01$*, and *** denotes* $p < .001$.

Results of the SEM show that there is a significant mediation effect of negative error response between MIL and ATA (b_{ab} = -.14, *SE* = .07, $p < .05$). There are significant effects between MIL and negative error response, MIL and conscientiousness, negative error response and ATA, self-efficacy and ATA, and conscientiousness and ATA. There is no significant mediation effect through self-efficacy and conscientiousness. Lastly, the SEM suggests there is a significant total effect of MIL on ATA (b_{total} = -.87, *SE* = .33, $p < .01$). Detailed results of the SEM are included in Appendix 8. Using the formula $f^2 = b^2 / (1 - b^2)$ and the standardized value for b_{std} (.127), Cohen's f^2 for the total effect was .02, indicating a small effect size (Cohen, 1988). The proportion of effect predicted by the mediated path from MIL to negative error response to ATA was only 16.09%, suggesting that there could be other mediating variables. This possibility was explored, and it was found that perceived family learning attitude accounted for a substantial proportion of mediated effect. Details will be presented in the discussion section.

Verifying Assumptions for T-Test and Regression Analysis

Assumptions for performing *t*-test and regression analysis between key variables were verified and presented in this section.

Negative Error Response vs. MIL

For the assumption of sample independence, since participants were recruited online, they came from varying cities. Therefore, sample data should be independent of each other. For the assumption of normality, Shapiro-Wilk normality test showed that the group with no MIL had $W = 0.966$ ($p < 0.05$), indicating normality assumption was violated. Bootstrap method will be used to check for confidence interval and verify significance in a later section. For the group with MIL, $W = .956$ ($p > 0.05$), indicating that normality assumption held. Regarding the assumption for homogeneity in variance, the difference in variance was $F = 0.759$ ($p > 0.05$); therefore, homogeneity in variance was acceptable. The mean and standard deviation for negative error response were 53.18 and 14.12. The box plot in Figure 7 shows that 2 samples are potential outliers in the non-MIL group (exceeding 3 standard deviations (95.55)). These outliers could unduly affect the measure and were eliminated from the analysis.

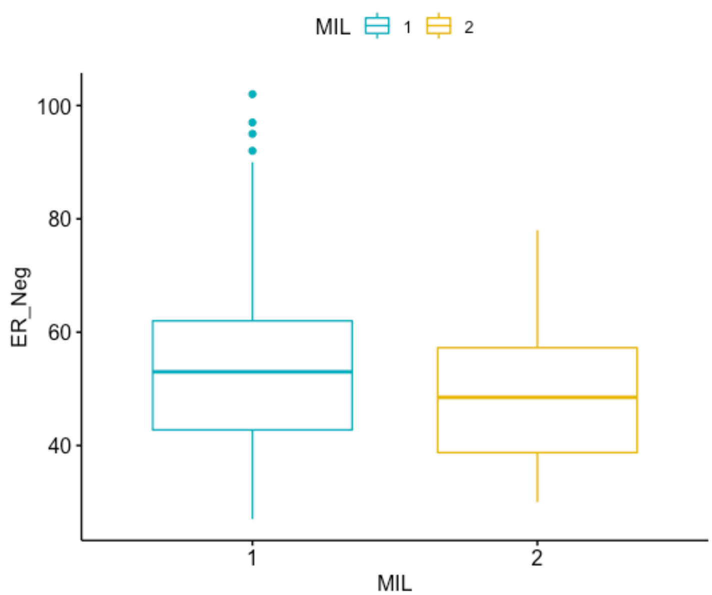

Figure 7. Box Plot for Negative Error Response by MIL Group.

Self-Efficacy vs MIL

For the measure of self-efficacy, regarding the assumption of normality, Shapiro-Wilk test values were $W = .977$ ($p = .442$) and $W = .992$ ($p = .107$) for MIL group and non-MIL group respectively, indicating that normality for both groups. Homogeneity in variance was $F = 1.02$, ($p = .872$), indicating that the assumption for homogeneity is also valid. The mean for self-efficacy is 37.75 and standard deviation is 11.23. The box plot in Figure 8 shows that one sample is a potential outlier in the MIL group that is 3 standard deviations below mean (4.05). This outlier could unduly affect the measure and was eliminated from the analysis.

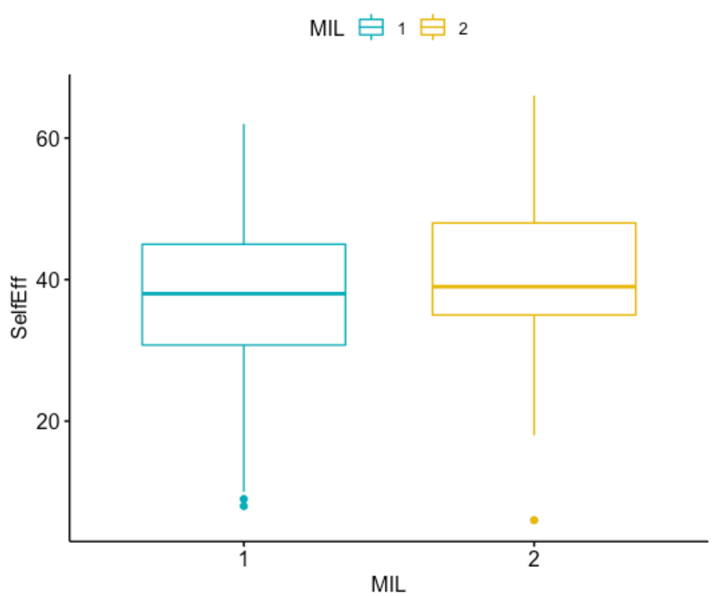

Figure 8. Box Plot for Self-Efficacy by MIL Group.

Conscientiousness vs MIL

For the measure of conscientiousness, regarding the assumption of normality, Shapiro-Wilk test values were $W = .946$ ($p = .02$) and $W = .984$ ($p = .003$) for MIL and non-MIL groups respectively, indicating that normality assumption is violated. Bootstrapping method will be used to test for confidence interval and verify significance in a later section. Homogeneity in variance is $F = .953$, ($p = .87$), indicating homogeneity in variance. The mean for conscientiousness was 31.94 and standard deviation was 6.70. The box plot in Figure 9 shows that one sample is a potential outlier in the MIL group that is 3 standard deviations below mean (11.83). This outlier could unduly affect the measure and was eliminated from the analysis.

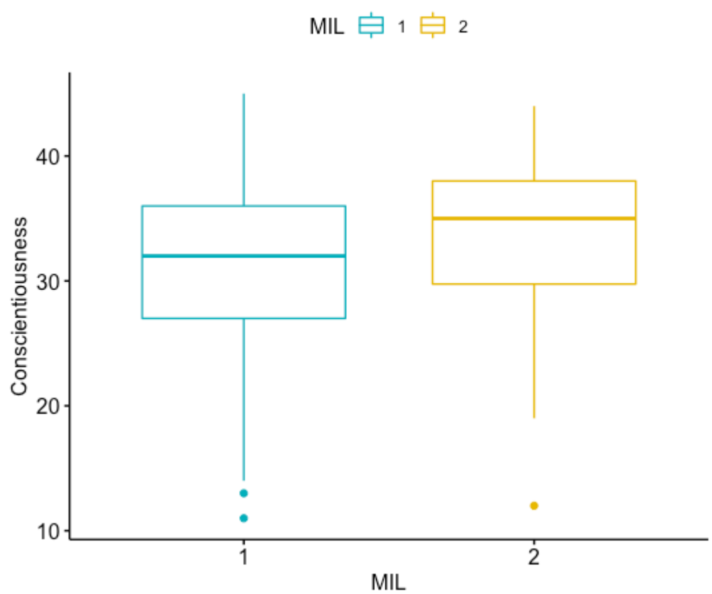

Figure 9. Box Plot for Conscientiousness by MIL Group.

ATA vs. Potential Mediating Variables

For the assumption of linearity, plot 2 in Figures 10-12 shows that most standardized residuals are within -2 to 2, indicating that linearity. Based on plots 1 and 3, no patterns were detected in residuals and square root of residuals. In addition, residuals were equally spread around the y = 0 line, indicating that the assumption for homoscedasticity was valid. Lastly, based on plot 2's Normal Q-Q plot, observations lie along the 45-degree line, indicating normality.

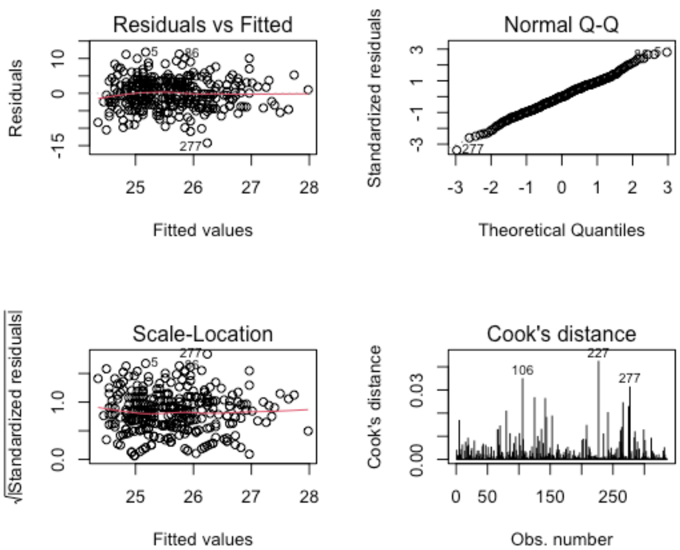

Figure 10. Residuals Plots for ATA vs Negative Error Response

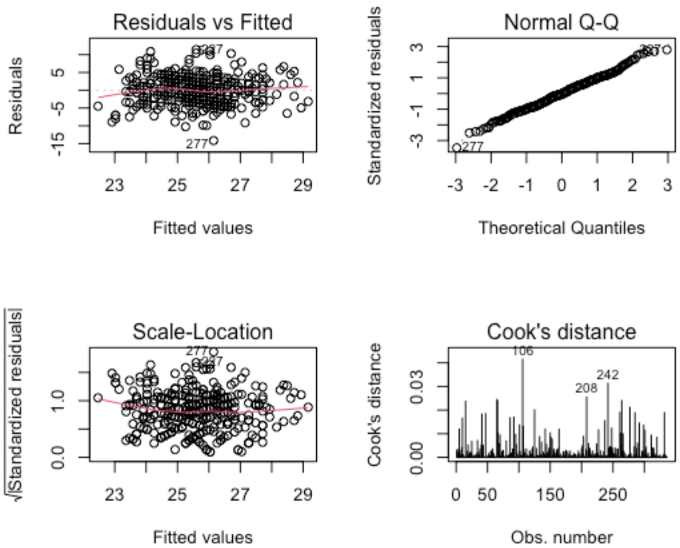

Figure 11. Residuals Plots for ATA vs. Self-Efficacy

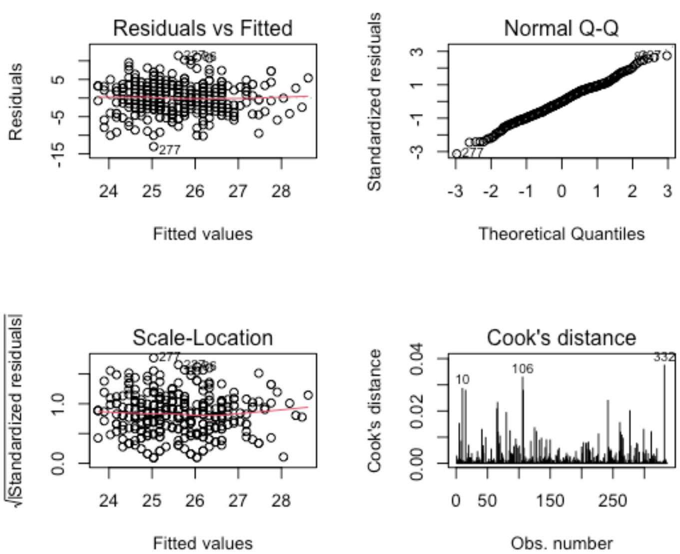

Figure 12. Residuals Plots for ATA vs. Conscientiousness

T-Tests and Linear Regression Analyses

T-tests in this section were conducted using the *t*.test program in the stats package in R. Welch two sample t-test was used to test the relationships between MIL and the potential mediating variables. Welch *t*-test was chosen because it is better than Student t-test for unequal sample sizes (between MIL and non-MIL groups) and unequal variances between two sample datasets.

Negative Error Response vs. MIL T-Test

The MIL group (N = 48) had a mean negative error response of 49.19 and standard deviation of 12.55. The non-MIL group (N = 288) had a mean negative error response of 53.85 and standard deviation of 14.28. The Welch two sample t-test resulted in $t(68.93) = 2.33$, $p = .02$. However, as stated in the previous section, normality assumption was not met $W = .97$ ($p < .05$). To correct for normality assumption violation, transformation of the data (such as log or square root transform) could be used. However, since transformation does not offer a very intuitive understanding of the relationship, non-parametric bootstrapping was chosen instead. The boot and boot.ci functions in the boot library (Canty & Ripley, 2022) in R with 5000 replicates were used. In addition, the boot.ci bias corrected and accelerated (BCa) option was used to correct for potential skew and bias in the bootstrapped data (DiCiccio & Efron, 1996). The estimated 95% confidence interval range for the t-value was .21 and 4.60, *95% CI* [.21, 4.60]. Since the CI range did not cross zero, one can reject the null hypothesis that MIL had no significant relationship with negative error response. Alternatively, it can be stated that MIL significantly and negatively predicted negative error response. Cohen's *d* was then calculated to gauge the effect size using the formula $d = \frac{2t}{\sqrt{df}}$. Cohen's *d* = .56, indicating a medium effect size between MIL and negative error response (Cohen, 1988).

To control for potential confounds, multiple regression was used for the analysis (Regression with Categorical Variables, n.d.). Referencing prior studies (Guhn et al., 2020; Hille & Schupp, 2015), when controlling for maternal and fraternal education, maternal and fraternal ethnicity, age, gender, and grade, the regression coefficient for MIL remained significant $b = -4.87$, $t(327) = -2.14$, $p < .05$. To overcome normality issue, bootstrapping method was applied to the multiple regression analysis and resulted in the confidence interval for the regression coefficient *95% CI* [-9.18, -.44]. The R code for bootstrapping and multiple regression is provided in Appendix 9. Since the confidence interval did not cross zero, it can be concluded that MIL had a significant effect on negative error response even while controlling for the potential confounding variables.

Self-Efficacy vs. MIL T-Test

The MIL group ($N = 48$) had a mean self-efficacy of 39.56 and a standard deviation of 11.46. The non-MIL group ($N = 288$) had a mean negative error response of 37.45 and a standard deviation of 11.19. The Welch two sample t-test resulted in $t(62.85) = -1.19$, $p = .24$, indicating that the association between MIL and self-efficacy was not significant.

Conscientiousness vs. MIL T-Test

The MIL group ($N = 48$) had a mean conscientiousness of 33.67 and a standard deviation of 6.49. The non-MIL group ($N = 288$) had a mean conscientiousness of 31.65 and standard deviation of 6.71. The Welch two sample t-test resulted in $t(64.89) = 1.98$, $p = .05$. However, as stated in the previous section, normality assumptions were not met. Shapiro-Wilk tests were $W = .95$ ($p < .05$) and $W = .984$ ($p < .01$) for MIL and non-MIL group respectively. Again, using the non-parametric bootstrapping method with 5000 replicates as above, the estimated 95% CI range for the *t*-value was .01 and 4.49, *95% CI* [.01, 4.49]. Since the CI range did not cross zero, one can reject the null hypothesis that MIL had no significant association with conscientiousness. Alternatively, it can be stated that MIL significantly predicted conscientiousness. Cohen's $d = .49$ indicated a medium effect size between MIL and conscientiousness (Cohen, 1988). However, when potential confounds were accounted for (maternal and fraternal education, maternal and fraternal ethnicity, age, gender, and grade), the regression coefficient for MIL became insignificant $b = 2.01$, $t(334) = 1.94$, $p = .054$. Using the same bootstrapping method, the resulting confidence interval also crossed zero, *95% CI* [-.23, 3.89], and thus it can be concluded that MIL's effect on conscientiousness was not significant when controlling for the confounding variables. CI results further suggested that certain effects were due to the confounding variables and not solely due to MIL.

ATA vs. Negative Error Response

Linear regression was used to investigate the association between ATA and negative error response. The function lm in the stats package in R was used. The linear regression results indicated that negative error response significantly predicted ATA, $b =$

.05, $t(334) = 2.96$, $p < .01$. The slope coefficient b matched the path coefficient from the mediation analysis using SEM. Negative error response also explained a significant proportion of variance in ATA, $R^2 = .03$, $F(1, 334) = 8.77$, $p < .01$. Effect size was further evaluated using Cohen's f^2 according to the following formula $f^2 = \frac{R^2}{1-R^2}$ (Cohen, 1988). $f^2 = .03$ indicating a small effect size between negative error response and ATA. According to Cohen's (1988) guideline, $f^2 \geqslant 0.02$, $f^2 \geqslant 0.15$, *and* $f^2 \geqslant 0.35$ represent small, medium, and large effect sizes, respectively.

To control for the effects of potential confounds, multiple regression was used to test the relationship between negative error response and ATA while holding other variables constant. This was done using multiple regression via the lm function in R. The variables maternal and fraternal education, maternal and fraternal ethnicity, gender, grade, age were held constant. Parents' education has been found to be associated with parenting style (Kashahu et al., 2014), which could in turn affect negative error response and ATA. While controlling for the potential confounds, the effect of negative error response on ATA remained significant, $b = .03$, $t(327) = 3.32$, $p < .01$. After controlling for demographics, I further expanded the list of controlled variables to include conscientiousness and academic performance. Although conscientiousness is hypothesized as a potential mediating variable, given its correlation with MIL and ATA in prior research (Hille & Schupp, 2015; Lay & Brokenshire, 1997), conscientiousness also has the possibility of being a confound that affects both negative error response and ATA. Similarly, academic performance has been found to be associated with MIL (Guhn et al., 2020) and indirectly with ATA through procrastination (Milgram et al., 1995), and thus, I also tested their roles as potential confounds. When adding conscientiousness and academic performance into the multiple regression analysis, as level 2 and 3 variables respectively, negative error response and ATA maintained a significant association, *b* = .03, *t*(325) = 2.99, $p < .01$.

ATA vs. Self-Efficacy

The results of linear regression indicated that self-efficacy significantly predicted ATA, *b* = -.11, *t*(334) = -5.61, $p < .001$. The slope coefficient b matched that from the mediation analysis using SEM. Self-efficacy also explained a significant proportion of variance in ATA, $R^2 = .09$, $F(1,334) = 31.44$, $p < .01$. Cohen's $f^2 = .10$ indicating a small effect size between self-efficacy and ATA. When controlling for demographics, maternal and fraternal education, maternal and fraternal ethnicity, gender, grade, and age, the effect of self-efficacy on ATA was reduced but remained significant, *b* = -.05, *t*(327) = -4.53, $p < .001$.

ATA vs. Conscientiousness

The results of linear regression indicated that conscientiousness significantly predicted ATA, *b* = -.14, *t*(334) = -4.22, $p < .001$. The slope coefficient b also matched that from the mediation analysis using SEM. Conscientiousness explained a significant proportion of variance in ATA, $R^2 = .05$, $F(1,334) = 17.80$, $p < .01$. Cohen's $f^2 = .05$ indicated a small effect size between conscientiousness and ATA. When controlling for the same potential confounds, the effect of conscientiousness on ATA was reduced but remained significant, *b* = -.06, *t*(327) = -3.09, $p < .01$.

Testing for V-Structures

V-structures are composed of 3 variables, for example, X, Y, and Z, where X and Z are correlated, Y and Z are correlated, but X and Y are not correlated. However, when X and Y become correlated while controlling for Z, it implies that X has a causal effect on Z and Y has a causal effect on Z (Pearl, 2016). Based on the DAG in Figure 5, it was shown that self-efficacy and opportunity cost have effects on ATA. The hypothesis for a v-structure was that self-efficacy and opportunity cost were not correlated but each was individually correlated to ATA. In addition, self-efficacy and opportunity cost would become correlated when conditional on ATA. Using cor.test function in stats package in R, there was significant correlation between self-efficacy and opportunity cost even when not controlling for ATA, $r(334) = -.29$, $p < .001$. Hence, the hypothesized v-structure could not be verified.

Other potential v-structures were then tested based on the literature search of the correlates of ATA and the correlation matrix in Table 2. Partial correlations were calculated using the pcor.test function in the ppcor package in R. The results are displayed in Table 4.

Table 4. Bivariate and Partial Correlations.

X	Y	Z	*r*(X,Y)	*r*(X,Z)	*r*(Y,Z)	*r*(X,Y | Z)
Positive Error Response	MIL	ATA	-.04 p = .48	-.28 p < .001	-.13 p = .02	-.08 p = .16
Perceived Opportunity Cost	MIL	ATA	-.05 p = .39	.38 p < .001	-.13 p = .02	3.5e-4 p = .99
Perceived Autonomy	MIL	ATA	8.8e-3 p = .87	-.37 p < .001	-.13 p = .02	-.04 p = .46
Family Learning Attitude	Negative Error Response	ATA	9.99e-3 p = .86	-.24 p < .001	.19 p < .001	.06 p = .28

Verifying potential v-structures using bivariate and partial correlations.

As the results in Table 4 indicate, the partial correlations were not significant while controlling for the third variable. Therefore, no v-structure could be established, and causal effects cannot be verified. However, partial correlations indicated a substantial increase in association when controlling for the collider variable. There were significant increases in correlation as well as reductions in p-value when controlling for the collider variables for some of the tested v-structures.

Chapter IV.

Discussion

The present study contributes to existing research on academic task aversiveness (ATA) by examining its relations with musical instrument learning (MIL), self-efficacy, conscientiousness, and negative error response. There was no prior study investigating the relationship between MIL and ATA. The analyses in this study suggested that the association between MIL and ATA was mediated through self-efficacy (in the Bayesian network analysis) and negative error response (in the SEM). This study also confirmed certain correlates of MIL and ATA that have been suggested by prior studies. In this section, the hypotheses that this study set out to test will be reviewed, followed by a discussion on the differences between the results of the Bayesian network and SEM. One interesting finding that was not part of the original hypotheses – but was supported by the data collected – will also be presented: mediation through perceived family learning attitude. This will be followed by a summary of limitations and suggestions for possible future work.

Review of Hypotheses

This study was developed based on the theory of transfer. Based on the analysis, transfer does take place from the area of musical instrument learning (MIL) to the area of formal education. This study set out to test the hypothesis that MIL is associated with academic task aversiveness (ATA). It was hypothesized that the effects of MIL would be mediated through three variables, namely negative error response, self-efficacy, and conscientiousness (Figure 13).

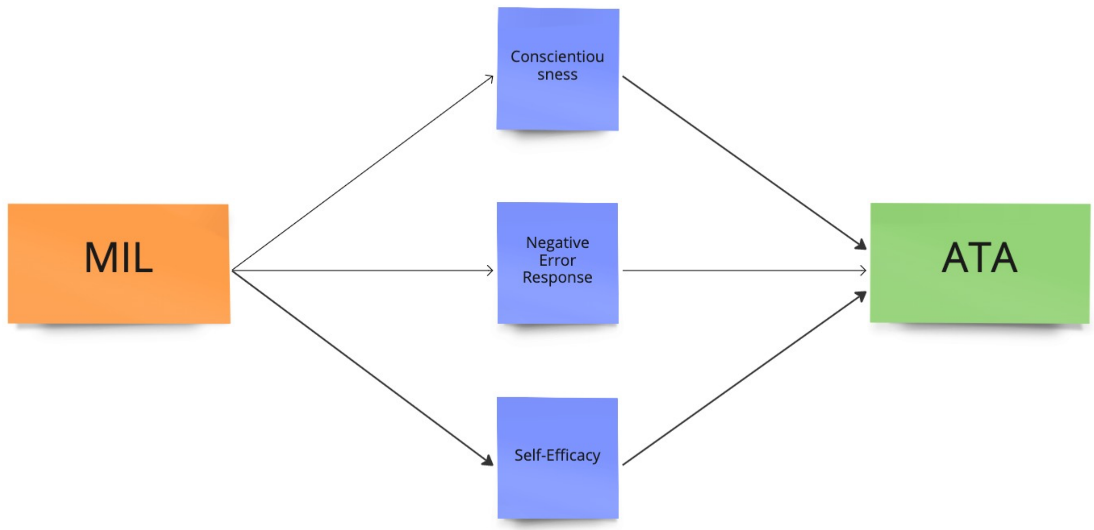

Figure 13. Conceptual Diagram.

Before testing the hypotheses, when checking for internal consistencies of the measures, it was found that ATA had a low Cronbach alpha. When I applied exploratory factor analysis on the data, I found that a 3-factor model offered the best fit – aversiveness for high cognitive demand tasks, low cognitive demand tasks, and memorization tasks. The data for high cognitive demand tasks were then used for the remainder of the analyses.

Three methods were used to analyze the data: Bayesian network analysis, structural equation modeling (SEM), and regression analysis. After limiting years of musical instrument learning to 8 to 11 years, Bayesian network analysis showed weak mediating path from MIL to ATA through self-efficacy, whereas SEM showed a significant mediating path from MIL to ATA through negative error response. Regression analyses confirmed the results of SEM even after accounting for normality assumption violations.

Verifying Assumptions and Accounting for Potential Confounds

After using SEM to understand the mediating effects of the variables, traditional regression methods were used to verify the results. Regression methods were chosen because they allowed for the use of bootstrapping to overcome normality violation issues and multiple regression to control for potential confounds. When checking for violations of the assumptions for regression, it was found that normality assumptions were not satisfied for negative error response vs. MIL and conscientiousness vs. MIL. In the case of negative error response, applying bootstrapping method on multiple regression to control for potential confounds resulted in a regression coefficient that did not cross zero, *95% CI* [-9.18, -.44]. Since potential confounds were accounted for, it allowed for stronger causal inferencing that MIL has a significant and causal effect on negative error response.

However, in the case of conscientiousness vs. MIL, when applying bootstrapping on multiple regression to control for potential confounds, the confidence interval for the regression coefficient crossed zero, *95% CI* [-.23, 3.89]. That implied some of the effects between MIL and conscientiousness are in part due to the potential confounds, and when those are accounted for, the effect due to MIL is no longer significant.

In terms of the effects on ATA, it was found that negative error response (b = .03, $t(327) = 3.32$, $p < .01$), self-efficacy (b = -.05, $t(327) = -4.53$, $p < .001$), and conscientiousness (b = -.06, t(327) = -3.09, p < .01) had significant correlations with ATA while holding demographics variables constant, confirming the results of the SEM.

Differences between Bayesian Network Analysis and SEM

Ideally, when all the assumptions are met, Bayesian networks show the causal effects of a system. These assumptions include the existence of a DAG underlying the network, the lack of unobserved variables (causal sufficiency), the causal Markov assumption (independence), and dependencies between unlinked variables (causal faithfulness) (Briganti et al., 2022; McNally, 2016). In this study, certain correlates of ATA – intelligence, executive functions, and personality – have not been included due to the difficulty in obtaining such data. In addition, the causal faithfulness assumption or dependencies between variables could not be verified. Therefore, causal effects cannot be concluded from the analysis. However, Bayesian networks can suggest potential causal relations and probabilistic dependencies of the variables (McNally, 2016; McNally et al., 2022). In this case, it shows that higher ATA more strongly predicts higher self-efficacy and in turn, MIL, rather than the other way around.

When comparing the Bayesian network and SEM, results indicated that the models did not show the same mediating path. The Bayesian network showed a mediating pathway through self-efficacy, whereas SEM showed a mediating pathway through negative error response. One possible reason is that there could be bi-directional relationships in the network. When that is the case, Bayesian networks would not include the path in the DAG. In this study, ATA and negative error response may be inter-dependent – stronger error response could lead to high aversiveness for academic tasks and vice versa. Another possible reason is that there could be cyclic relationships, which are not allowed in the DAG. When cyclic relationships exist in the Bayesian network, the paths that lead to cyclic relationships are eliminated from the DAG. In this instance, the boot.strength function in bnlearn produced the warning message that the arc from negative error response to MIL was eliminated because the arc introduced a cycle in the DAG. Therefore, due to these factors, it is reasonable to observe different results between Bayesian network analysis and SEM.

Upon closer examination of the SEM results, I found several aspects to be curious. Firstly, the direct effect of MIL on conscientiousness and direct effect of conscientiousness on ATA are both significant. However, the indirect effect of MIL on ATA through conscientiousness was not. Participants who have MIL are more likely to have higher conscientiousness, and separately, participants with higher conscientiousness are more likely to have lower ATA. However, participants who have higher conscientiousness and MIL were not significantly likely to have lower ATA. One interpretation is that contrary to my original hypothesis, participants who have higher conscientiousness due to MIL actually causes them to have higher ATA. One study (Charalambous et al., 2019) showed similar relations in its mediation analysis.

The second curious observation was that two of the indirect effects (through conscientiousness and self-efficacy) and the direct effect of MIL on ATA were not significant. However, the total effect of the model was significant. One interpretation is that there are other mediating variables that have not been included in the model, and majority of the effect of MIL on ATA is acting through the missing variables.

Perceived Family Learning Attitude

With the possibility of other mediating variables, I set out to test the indirect effects of other variables measured in the survey. I started with those that showed higher correlations with MIL – positive error response, growth mindset, academic performance, and perceived opportunity cost. However, none of the indirect effects were significant. Lastly, I tested perceived family learning attitude (Appendix 10) and found that the indirect effect was indeed significant.

The perceived family learning attitude scale was created for this study because, after a brief search of the literature, it was found that there has not been a scale created to measure families' and parents' attitudes toward learning. Since this was a survey completed by the children and gauged by the children, it measured the subjectively perceived attitude by the children regarding their parents and family. Although there have not been studies between family learning attitudes and ATA, prior studies support the notion that parents' attitudes could have an impact on students' perception of schoolwork.

Multiple studies have investigated parents' attitudes toward learning and education. One particular study by Marjoribanks (1987) investigated parents' attitudes toward children's schoolwork, independence, and expected education level. It was found that parents' overall attitude toward learning was associated with children's academic performance as well as their attitude toward school. Children of parents with better attitude ratings scored higher on items such as "I like being in this school", "I enjoy reading", and "Doing well in school is most important to me". Another study suggested that parents' own attitudes toward science were associated with children's science achievement (Perera, 2014). The scale for parents' attitudes toward science included statements such as "Advances in broad science and technology usually improve people's living conditions" and "There are opportunities for me to use broad science in my everyday life". The study considered potential confounding variables such as grade, gender, and immigrant status. Although not directly measuring children's attitudes toward schoolwork (such as ATA), these studies suggested that parents' attitudes toward school and education could have an effect on children's views of school as well as performance. Both studies referenced above were based on survey results of parents. However, children's perception of their parents' attitudes should have an equal or stronger impact on the children (Jourard & Remy, 1955).

In this study, mediation analysis showed a significant indirect effect from MIL to ATA through family learning attitude, $b = -.33$, $p < .001$. The effect of MIL on family learning attitude was $b = 1.39$, $p < .001$, and the effect of family learning attitude on ATA was $b = -.24$, $p < .001$. Due to the multiple number of hypotheses tested, familywise error rates (FWER) had to be accounted for. The Holm-Bonferroni Method (also called Holm's Sequential Bonferroni Procedure; Shaffer, 1986) was used to reduce the possibility of obtaining an incorrect statistically significant result (i.e., a Type I error). The formula used was HB = Target Alpha Level / (n – significance rank number + 1). The ranking of the p-values of the hypotheses tested is as follows:

Family Learning Attitude $b = -.33$, $p = 2e\text{-}16$

Academic Performance $b = -.052$, $p = .36$

Perceived Opportunity Cost $b = -.121$, $p = .37$

Positive Error Response $b = .075$, $p = .454$

Growth Mindset $b = 0.0021$, $p = .94$

Calculating *HB* for family learning attitude $HB = 0.05 / (5 – 1 +1) = .01$. Comparing family learning attitude's *p-value* (2e-16) to *HB* (.01), the *p-value* was less than *HB*. Therefore, the null hypothesis could be rejected, and one could conclude that the mediating effect of family learning attitude was significant. Notably, the p-value ranking would not change even when including the original hypotheses (self-efficacy, negative error response, and conscientiousness), as family learning attitude had the smallest p-value among those as well.

When family learning attitude was added to the SEM analysis, the results were as shown in Figure 14. The mediating effects of negative error response (-.14, $p = .048$) and family learning attitude (-.33, $p = .003$) together ($b_{ab} + b_{ij} = -.47$) account for 54% (compared to 16.1% for negative error response alone) of the total effect (-.87) between MIL and ATA.

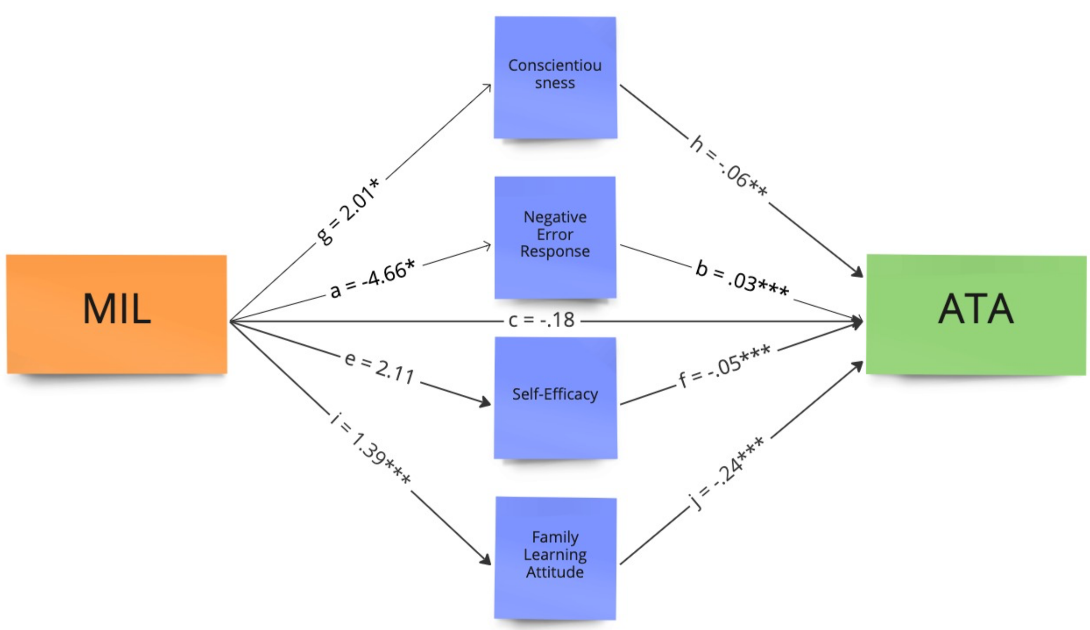

Figure 14. SEM Path Diagram with Family Learning Attitude.

Path coefficients using structural equation modeling with the Lavaan package in R. Indirect effects: $b_{ab} = -.14, p < .05; b_{ef} = -.11, p = .26; b_{gh} = -.11, p = .14; b_{ij} = -.33, p < .01$*. Total effect* $b_c + b_{ab} + b_{ef} + b_{gh} + b_{ij} = -.87, p < .01$*. * denotes* $p < .05$*, ** denotes* $p < .01$*, and *** denotes* $p < .001$*.*

Limitations

Due to the COVID-19 lockdown situation, the high school population could not be accessed at the time of the study. As a result, only a small number of participants who are currently engaging in MIL could be recruited through crowdsourcing websites. After limiting the years of training to 8 to 11 years, the number was further reduced to 48. As shown by the effect sizes and power analysis, the power of the regression analysis required a higher number of participants (approximately 90). As such, the significances (p-values) of the analyses would be more accurate if the number of participants was higher.

Certain potential confounding variables were not measured and controlled for in this study since collecting those data would have substantially lengthened the survey. Variables such as openness to new experiences (Hille & Schupp, 2015), SES, and parenting style have been shown to be associated with MIL (Corrigall & Schellenberg, 2015). These factors have been shown to correlate with academic performance and could also influence ATA and be confounding variables.

In this study, the measure of perceived family learning attitude was shown to mediate the relationship between MIL and ATA. However, this measure was created for this study and have not been verified in terms of construct validity. Nonetheless the Cronbach alpha for this measure was .70, indicating an acceptable internal consistency.

Regarding representativeness and generalizability of the findings, the participants only included those with 8 to 11 years of MIL and were recruited from North America and the UK only. Notably, results might differ for Asian or other cultures as was suggested by a study on work avoidance (King & McInerney, 2014).

Future Research

Future research could benefit from a more focused age group with a higher proportion of participants who are still participating in MIL. The data collected showed that out of 875 participants, 148 have had more than 8 years of MIL. However, more than half of them (66) had stopped for more than 1 year. A younger age group would offer a higher percentage of participants who are still participating in MIL. Since the crowdsourcing websites that were surveyed in this study do not offer participants who are underaged, other channels would be required to recruit younger participants.

In the present study, the survey did not collect information on the number of hours participants practice per week for those who are not currently engaged in MIL. According to several studies, the number hours of practice per week is associated with various aspects, such as musical achievement (Sloboda et al., 1996) and executive function performance (Loui et al., 2019). It would be meaningful to examine the relationship between practice duration and the correlates of this study and also whether there is a threshold for the minimum number of hours that students of MIL need to dedicate in order to benefit from it.

Due to the practical limit on the size of online surveys (surveys that are too long tend to receive inaccurate results) (Palmer & Strickland, 2016), and the non-physical nature of the medium, many factors were omitted from this study. Factors that could mediate the relationship between MIL and ATA and were not able to be collected in this study include IQ, executive function performances, socioemotional functioning, and audiation. Future studies that take place in a campus environment rather than crowdsourcing surveys could be beneficial as those lengthier and physical aspects could be measured.

Future studies could utilize more sophisticated modeling methods for SEM. Multiple studies have demonstrated that various features of MIL, such as the years of musical training (Guhn et al., 2020), hours of practice (Sloboda et al., 1996), level achieved, private vs. group training (Holochwost et al., 2021), and years lapsed could have an impact on the dependent variables. As such, modeling MIL as a latent variable could provide more revealing results than a categorical MIL variable.

Appendix 1.

Musical Instrument Learning Questionnaire

- Are you currently in a school band or orchestra?
    - Yes
        - Which one?
        - On average, how many hours do you practice a week for that instrument?
        - When did you start playing that instrument? (year)
    - No
        - Have you participated in a school band or orchestra before?
            - Yes
                - Which one?
                - When did you stop? (year)
                - When did you start? (year)
            - No
- Are you currently taking any private music lessons?
    - Yes
        - What instrument or type of music lessons?
        - On average, how many hours do you practice a week for that instrument?
        - When did you start playing that instrument? (year)
        - Have you taken any music exams?
            - Yes
                - What was the highest level exam you passed on?
                - What board was the exam from? (choose one of the following)
                    - ABRSM (Level 1-8, DIPASRSM, LRSM, FRSM, LRAM, ARCM)
                    - MTNA
                    - MTAC
                    - Other
            - No

        - What is the approximate level of music you play?
        - What exam board is the level approximation based on? (choose one of the following)
          - ABRSM (Level 1-8, DIPASRSM, LRSM, FRSM, LRAM, ARCM)
          - MTNA
          - MTAC
          - Other
- No
  - Have you taken private music lessons before?
    - Yes
      - What type of private music lessons?
      - When did you stop? (year)
      - When did you start? (year)
    - No

Appendix 2.

Error Response Instruction and Scenarios

We'd like to know how you usually react when you realize you've just made a mistake. There are 5 scenarios. After reading each scenario, please let us know how you feel by rating the scales.

Scenario 1: “You just spent 10 minutes working on a math problem. Now, you realized you made a mistake in the first few steps and have to redo the whole solution again”.

Scenario 2: “You are preparing for a history test. You have gone over some key information like historical figures and dates a couple of times. You are quizzing yourself and found that you just remembered the same historical figure wrongly for the third time.”

Scenario 3: “You worked really hard over the weekend to complete an assignment that you thought was due on Monday. Then when you came to class, you realized that the assignment is actually due a week later.”

Scenario 4: “You are collecting data for a science experiment. You have spent an hour on it and are half-way through. Now you realize that you didn’t set up the equipment properly and the measurements are therefore not accurate. You have to correct the setup and start the experiment again.”

Scenario 5: “You are writing a 5-page essay for English literature/language arts class. You are almost finished and are writing the ending section. Now you realize a couple of the previous sections aren't coherent and have to rewrite some portions of the review. This is going to take you another 2 hours.”

Appendix 3.

I-PANAS-SF

1 = Very slightly/not at all; 2 = A little; 3 = Moderately; 4 = Quite a bit; 5 = Extremely

Upset

Hostile

Alert

Ashamed

Inspired

Nervous

Determined

Attentive

Afraid

Active

Appendix 4.

General Measure of Self-Efficacy

(Remarks in brackets below do not appear in the actual survey. Items marked with R are reverse-scored items.)

Instructions: The following statements describe people's feelings and reactions to various situations. Please read each statement carefully and describe the extent to which you agree with each statement, using a 5-point scale where 1 indicates "strongly disagree" and 5 indicates "strongly agree".

Answer categories

1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree;

1. When I make plans, I am certain I can make them work.
2. One of my problems is that I cannot get down to work when I should. (R)
3. If I can't do a job the first time I keep trying until I can.
4. When I set important goals for myself, I rarely achieve them. (R)
5. I give up on things before completing them. (R)
6. I avoid facing difficulties. (R)
7. If something looks too complicated, I will not even bother to try it. (R)
8. When I have something unpleasant to do, I stick to it until I finish it.
9. When I decide to do something new, I go right to work on it.

10. When trying to learn something new, I soon give up if I am not initially successful. (R)

11. When unexpected problems occur, I don't handle them well. (R)

12. I avoid trying to learn new things when they look too difficult for me. (R)

13. Failure just makes me try harder.

14. I feel insecure about my ability to do things. (R)

15. I am a self-reliant person.

16. I give up easily. (R)

17. I do not seem capable of dealing with most problems that come up in life. (R)

Appendix 5.

Big Five Inventory – Conscientiousness Scale

(Remarks in brackets below do not appear in the actual survey. Items marked with Consc are part of the conscientiousness scale. Items marked with Consc R are reverse-scored items.)

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

1 = Disagree Strongly
2 = Disagree a little
3 = Neither agree nor disagree
4 = Agree a little
5 = Agree strongly

I am someone who…

1. does a thorough job (Consc)
2. is original, comes up with new ideas
3. can be somewhat careless (Consc R)
4. is relaxed, handles stress well
5. is a reliable worker (Consc)
6. has a forgiving nature
7. tends to be disorganized (Consc R)

8. tends to be lazy (Consc R)
9. has an assertive personality
10. perseveres until the task is finished (Consc)
11. is sometimes shy, inhibited
12. does things efficiently (Consc)
13. prefers work that is routine
14. makes plans and follows through with them (Consc)
15. is easily distracted (Consc R)

Appendix 6.

Academic Task Aversiveness Questionnaire

For each of these tasks, imagine that it is due soon, and you need to be working on it. Please rate the degree to which you find the task enjoyable. 1 being very unpleasant; 5 being very enjoyable.

Answer categories

1 = Very Unpleasant; 2 = Unpleasant; 3 = Neutral; 4 = Enjoyable; 5 = Very Enjoyable

1. writing a term paper
2. studying for an exam
3. keeping up with weekly reading assignments
4. performing administrative tasks, e.g., registering for courses or buying school supplies
5. attending meetings for team projects
6. practicing math problems
7. memorizing historical facts
8. learning new physics concepts

Appendix 7.

Bnlearn Code

```
# A tutorial on Bayesian networks for psychopathology researchers
# Briganti G, Scutari M, McNally RJ

rm(list = ls()) # clear the workspace

# install.packages("bootnet")
# install.packages("ggplot2")

library(stats)
library(qgraph)
library(readr)
library(bootnet)
library(dplyr)
library(corpcor)
library(bnlearn)
library(psych)
library(ggplot2)


setwd("~/Library/Mobile
Documents/com~apple~CloudDocs/Documents/R_home/briganti")

NumBoots = 1000 # Sets the number of boots in all bootstrap calc

data <- read_csv("Data Stop 0_1 MIL 8_11a_MIL_ATAhia.csv",
        trim_ws = TRUE)

names <- c("Age",
    "FamLearn",
    "AcadPerf",
    "MIL",
    "ER_Neg",
    "ER_Pos",
    "Growth",
    "ATA",
    "OpCost",
    "Auton",
    "SelfEff",
```

```
          "Consc",
          "SportsPerf",
          "FratEth",
          "FratEd",
          "MatEth",
          "MatEd")

longnames <- c("Age",
           "Perceived Family Learning Attitude",
           "Academic Performance",
           "Musical Instrument Learning",
           "Error Response Negative Affect",
           "Error Response Positive Affect",
           "Growth Mindset",
           "Academic Task Aversiveness",
           "Perceived Opportunity Cost",
           "Perceived Autonomy",
           "Self-Efficacy",
           "Conscientiousness",
           "Sports or Performance Arts",
           "Fraternal Ethnicity",
           "Fraternal Education",
           "Maternal Ethnicity",
           "Maternal Education")

data <- as.data.frame(matrix(as.numeric(as.matrix(data)),
                      ncol=ncol(data),
                      byrow=TRUE))  #set data as numerics

bklist <- matrix(c("V1","V16", "V2","V16", "V3","V16",
              "V4","V16", "V5","V16","V6", "V16",
              "V7","V16", "V8","V16", "V9", "V16",
              "V10","V16", "V11","V16","V12","V16",
              "V13","V16", "V14","V16", "V15","V16",
              "V17","V16",
              "V1","V17", "V2","V17", "V3","V17",
              "V4","V17", "V5","V17","V6", "V17",
              "V7","V17", "V8","V17", "V9", "V17",
              "V10","V17", "V11","V17","V12","V17",
              "V13","V17", "V14","V17", "V15","V17",
              "V16","V17",
              "V1","V14", "V2","V14", "V3","V14",
              "V4","V14", "V5","V14","V6", "V14",
              "V7","V14", "V8","V14", "V9", "V14",
              "V10","V14", "V11","V14","V12","V14",
              "V13","V14", "V15","V14", "V16","V14",
```

```
                    "V17","V14",
                    "V1","V15", "V2","V15", "V3","V15",
                    "V4","V15", "V5","V15","V6", "V15",
                    "V7","V15", "V8","V15", "V9", "V15",
                    "V10","V15", "V11","V15","V12","V15",
                    "V13","V15", "V14","V15", "V16","V15",
                    "V17","V15"
                   ),
                  byrow = TRUE, ncol = 2)

    wtlist <- matrix(c("V4","V7", "V4","V11", # MIL to Growth, MIL to Self-
efficacy
                    "V4","V3", "V3","V4", # MIL and Academic Performance
                    "V4","V5", "V5","V4", # MIL and Error Response (neg scale)
                    "V4","V12", "V12","V4", # MIL and Conscientiousness
                    "V8","V3", "V3","V8", # ATA and Academic Performance
                    "V8","V11", "V11","V8", # ATA and Self-Efficacy
                    "V8","V10", "V10","V8", # ATA and Perceived Autonomy
                    "V8","V9", "V9","V8", # ATA and Perceived Opportunity Cost
                    "V11","V13", "V13","V11", # Self-efficacy and Sports/Performance
art participation
                    "V11","V7", "V7","V11", # Growth mindset and Self-efficacy
                    "V11","V5", "V5","V11", # Error response (neg scale) and Self-
efficacy
                    "V5","V7", "V7","V5"), # Growth mindset and Error response (neg
scale)
                  byrow=TRUE, ncol = 2)

    wtlistuni <- matrix(c("V4","V7", "V4","V11"), # MIL to Growth, MIL to Self-
efficacy
                   byrow=TRUE, ncol = 2)

    #### Create a matrix with V1 to V7 in column 1 and label names in column 2.
First create a vector "labeldata" with all the info by column:
    labeldata<-c("V1","V2","V3","V4","V5","V6","V7","V8","V9","V10",
            "V11","V12","V13","V14","V15","V16","V17",
            "Age","FamLearn","AcadPerf","MIL","ER_Neg","ER_Pos",
            "Growth","ATA","OpCost","Auton","SelfEff","Consc",
            "SportPerf","FratEth","FratEdu","MatEth","MatEdu")
    #### Then create the matrix "namemat" using the data in that vector
    namemat<-matrix(labeldata,length(names),ncol=2)

    # PC algorithm

    BNpc<-pc.stable(data, blacklist=bklist, whitelist = wtlist)
    pdf("pcdag.pdf", width=13, height=6)
```

```
qgraph(BNpc, labels=names,
      nodeNames=longnames,
      legend.cex=.5,
      vsize=6
      #,layout="circle")
      )
dev.off()

# Boot stability for the DAG
# Strength: connection strength, e.g. 0.85 >
#            connection appears in 86% of the fitted networks.
# Direction: probability of the direction
#     e.g. 0.57 means that in 57% of the fitted networks the connection goes in
#             the direction depicted in the graph.

BST <- boot.strength(data,
                 R = NumBoots,
                 algorithm = "pc.stable",
                 algorithm.args=list(blacklist=bklist, whitelist = wtlist)  )
head(BST)
qgraph(BST) # visualize output with qgraph
BST[BST$strength > 0.85 & BST$direction > 0.5, ]

avgnet1 <- averaged.network(BST,
                      threshold = 0.85)
avgnet1

bnlearn::score(avgnet1, data = data)

astr1 <- arc.strength(avgnet1,
                 data, "bic-g")  # compute edge strengths

#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)

astr1nrow<-nrow(astr1wmat)
```

Take the row names in the weight matrix and put it into a new matrix called "astr1rown" for merging later on.

```
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)
```

Create a matrix with V1 to V7 in column 1 and label names in column 2. First create a vector "labeldata" with all the info by column:

```
labeldata<-c("V1","V2","V3","V4","V5","V6","V7","V8","V9","V10",
          "V11","V12","V13","V14","V15","V16","V17",
          "Age","FamLearn","AcadPerf","MIL","ER_Neg","ER_Pos",
```

```
            "Growth","ATA","OpCost","Auton","SelfEff","Consc",
            "SportPerf","FratEth","FratEdu","MatEth","MatEdu")
#### Then create the matrix "namemat" using the data in that vector
namemat<-matrix(labeldata,length(names),ncol=2)

#### Merge astr1rown and namemat into new matrix "weightnamemat". This will
look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7

#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]

qgraph(astr1)
qgraph(astr1,labels=namevec)

pdf("DAGstable.pdf",
   width=13,
   height=6)
qgraph(astr1,
      labels=namevec,
      # nodeNames=longnames,
      legend.cex=.5,
      vsize=6
#      ,layout="circle")
      )
dev.off()


# Hill Climbing algorithm

# Boot stability for the DAG
# Strength: connection strength, e.g. 0.85 >
#           connection appears in 86% of the fitted networks.
# Direction: probability of the direction
#    e.g. 0.57 means that in 57% of the fitted networks the connection goes in
#             the direction depicted in the graph.

BST <- boot.strength(data, R = NumBoots,
                 algorithm = "hc",
                 algorithm.args = list(blacklist = bklist,whitelist = wtlistuni),
                 debug = TRUE)
head(BST)
BST[BST$strength > 0.85 & BST$direction > 0.5, ]
qgraph(BST)
```

```
avgnet1 <- averaged.network(BST,
                   threshold = 0.85)
avgnet1

bnlearn::score(avgnet1, data = data)

astr1 <- arc.strength(avgnet1, data, "bic-g")   ## compute edge strengths

#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)

astr1nrow<-nrow(astr1wmat)

#### Take the row names in the weight matrix and put it into a new matrix called
"astr1rown" for merging later on.
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)

#### Create a matrix with V1 to V7 in column 1 and label names in column 2.
First create a vector "labeldata" with all the info by column:
labeldata<-c("V1","V2","V3","V4","V5","V6","V7","V8","V9","V10",
          "V11","V12","V13","V14","V15","V16","V17",
          "Age","FamLearn","AcadPerf","MIL","ER_Neg","ER_Pos",
          "Growth","ATA","OpCost","Auton","SelfEff","Consc",
          "SportPerf","FratEth","FratEdu","MatEth","MatEdu")
#### Then create the matrix "namemat" using the data in that vector
namemat<-matrix(labeldata,length(names),ncol=2)

#### Merge astr1rown and namemat into new matrix "weightnamemat". This will
look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7

#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]

qgraph(astr1)
qgraph(astr1,labels=namevec)

pdf("hcDAGstable.pdf", width=13, height=6)
qgraph(astr1,
    labels=namevec,
    # nodeNames=longnames,
    legend.cex=.5,
    vsize=6
#      ,layout="circle")
)
```

```
dev.off()

#tabu algorithm
BNtabu <- tabu(data)
BST <- boot.strength(data, R = NumBoots,
        algorithm="tabu",
        algorithm.args = list(blacklist = bklist, whitelist = wtlistuni))
head(BST)
qgraph(BST) # visualize output with qgraph
BST[BST$strength > 0.85 & BST$direction > 0.5, ]
avgnet1 <- averaged.network(BST, threshold = 0.85)
avgnet1
bnlearn::score(avgnet1, data = data)
astr1 <- arc.strength(avgnet1,data, "bic-g")  # compute edge strengths
#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)

astr1nrow<-nrow(astr1wmat)

#### Take the row names in the weight matrix and put it into a new matrix called
"astr1rown" for merging later on.
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)

#### Create a matrix with V1 to V7 in column 1 and label names in column 2.
First create a vector "labeldata" with all the info by column:
labeldata<-c("V1","V2","V3","V4","V5","V6","V7","V8","V9","V10",
       "V11","V12","V13","V14","V15","V16","V17",
       "Age","FamLearn","AcadPerf","MIL","ER_Neg","ER_Pos",
       "Growth","ATA","OpCost","Auton","SelfEff","Consc",
       "SportPerf","FratEth","FratEdu","MatEth","MatEdu")
#### Then create the matrix "namemat" using the data in that vector
namemat<-matrix(labeldata,length(names),ncol=2)

#### Merge astr1rown and namemat into new matrix "weightnamemat". This will
look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7

#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]

qgraph(astr1)
qgraph(astr1,labels=namevec)
pdf("tabuDAGstable.pdf",width=13,height=6)
qgraph(astr1,
   labels=namevec,
```

```
    # nodeNames=longnames,
    legend.cex=.5,
    vsize=6
    #,layout="circle")
    )
dev.off()


# Equivalence classes

BNpc<-pc.stable(data,blacklist = bklist, whitelist = wtlist)
BNcp <- cpdag(BNpc)
pdf("cpdag.pdf", width=13, height=6)
qgraph(BNcp,
    labels=names,
    nodeNames=longnames,
    legend.cex=.5,
    vsize=6,
    layout="circle")
dev.off()

# vstructs(BNcp) # crashes

BST <- boot.strength(data, R = NumBoots,
            algorithm = "pc.stable",
            algorithm.args = list(blacklist = bklist, whitelist = wtlist),
            debug = TRUE,
            cpdag=TRUE)


head(BST)
BST[BST$strength > 0.85 & BST$direction > 0.5, ]
qgraph(BST)

avgnet1 <- averaged.network(BST,
                threshold = 0.85)
avgnet1

bnlearn::score(avgnet1, data = data)

astr1 <- arc.strength(avgnet1,
            data,
            "bic-g")   ## compute edge strengths

#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)
```

```
astr1nrow<-nrow(astr1wmat)
#### Take the row names in the weight matrix and put it into a new matrix called "astr1rown" for merging later on.
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)
#### Merge astr1rown and namemat into new matrix "weightnamemat". This will look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7
#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]
qgraph(astr1)
qgraph(astr1,labels=namevec)


pdf("cpDAGstable.pdf", width=13, height=6)
qgraph(astr1,
    labels=namevec,
    # nodeNames=longnames,
    legend.cex=.5,
    vsize=6
#      ,layout="circle")
)
dev.off()


# Grow shrink
BNgs <- gs(data,blacklist = bklist, whitelist = wtlist)
# BNgs <- gs(data,blacklist = bklist)
pdf("gsDAG.pdf", width=13, height = 6)
qgraph(BNgs,
    labels=names,
    nodeNames=longnames,
    legend.cex=.5,
    vsize=6,
    layout="circle")
dev.off()


BST <- boot.strength(data, R = NumBoots,
            algorithm = "gs",
            algorithm.args = list(blacklist = bklist, whitelist = wtlist),
            debug = TRUE)
# BST <- boot.strength(data, R = NumBoots,
#             algorithm = "gs",
#             algorithm.args = list(blacklist = bklist),
#             debug = TRUE)
```

```
head(BST)
BST[BST$strength > 0.85 & BST$direction > 0.5, ]
qgraph(BST)
write.csv(BST,"gs_bst.csv")

avgnet1 <- averaged.network(BST, threshold = 0.85)
avgnet1

bnlearn::score(avgnet1, data = data)

astr1 <- arc.strength(avgnet1, data, "bic-cg")   ## compute edge strengths, hybrid
astr1 <- arc.strength(avgnet1, data, "aic-cg")   ## compute edge strengths, hybrid
astr1 <- arc.strength(avgnet1, data, "loglik-cg")   ## compute edge strengths,
hybrid
astr1 <- arc.strength(avgnet1, data, "bic-g")   ## compute edge strengths, gaussian


#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)
astr1nrow<-nrow(astr1wmat)
#### Take the row names in the weight matrix and put it into a new matrix called
"astr1rown" for merging later on.
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)
#### Merge astr1rown and namemat into new matrix "weightnamemat". This will
look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7
#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]
qgraph(astr1)
qgraph(astr1,labels=namevec)

pdf("gsDAGstable.pdf", width=13, height=6)
qgraph(astr1,
     labels=namevec,
     # nodeNames=longnames,
     legend.cex=.5,
     vsize=6
#      ,layout="circle")
)
dev.off()



# Restricted Maximization hybrid algorithm
```

```
BNrs <- rsmax2(data,blacklist = bklist, whitelist = wtlistuni)
BST <- boot.strength(data,
              R = NumBoots,
              algorithm = "rsmax2",
              algorithm.args = list(blacklist = bklist, whitelist = wtlistuni),
              debug = TRUE)
head(BST)
BST[BST$strength > 0.85 & BST$direction > 0.5, ]

avgnet1 <- averaged.network(BST,
                threshold = 0.85)
avgnet1

bnlearn::score(avgnet1, data = data)

astr1 <- arc.strength(avgnet1, data, "bic-g")   ## compute edge strengths

#### Puts the weight matrix of "astr1" into a new matrix "astr1wmat"
astr1wmat <- getWmat(astr1)
astr1nrow<-nrow(astr1wmat)
#### Take the row names in the weight matrix and put it into a new matrix called
"astr1rown" for merging later on.
astr1rown <- matrix(rownames(astr1wmat),ncol=1,nrow=astr1nrow)
#### Merge astr1rown and namemat into new matrix "weightnamemat". This will
look up the correct name for each node.
weightnamemat<-merge(astr1rown,namemat,sort=FALSE)
#### If sort is TRUE, it will sort it back to default order from V1 to V7
#### Use column 2 of weightnamemat to create the name vector for labeling
namevec<-weightnamemat[1:nrow(weightnamemat),2]
qgraph(astr1)
qgraph(astr1,labels=namevec)

pdf("rsDAGstable.pdf", width=13, height=6)
qgraph(astr1,
    labels=namevec,
    # nodeNames=longnames,
    legend.cex=.5,
    vsize=6
#     layout="circle")
)
dev.off()


## GGM for comparison
n1 <- estimateNetwork(data, default="EBICglasso",
             threshold=TRUE)
```

```
pdf("glasso.pdf", width=13, height=6)
plot(n1,
#     layout="circle",
    labels=names,
    nodeNames=longnames,
    theme="colorblind",
    legend.cex=0.5,
    vsize=6)
dev.off()

plot(n1,
    #     layout="circle",

    nodeNames=longnames,
    theme="colorblind",
    legend.cex=0.5,
    vsize=6)
```

Appendix 8.

2nd SEM Detailed Results

```
> mediate<-read.csv("Data Stop 0_1 MIL8_11.csv",header=TRUE,sep=",")
> model4e<-'#parallel mediation
+ ER_Neg~a*MIL
+ ATAhi~b*ER_Neg
+ SelfEff~e*MIL
+ ATAhi~f*SelfEff
+ ATAhi~c*MIL
+ Conscientiousness~g*MIL
+ ATAhi~h*Conscientiousness
+ #indirect effect
+ ab:=a*b
+ ef:=e*f
+ gh:=g*h
+ #total effect
+ total:=c+(a*b)+(e*f)+(g*h)'
> fit4e<-sem(model4e,data=mediate,ordered=c("MIL"))
Warning messages:
1: In lav_data_full(data = data, group = group, cluster = cluster,  :
  lavaan WARNING: exogenous variable(s) declared as ordered in data: MIL
2: In lav_partable_check(lavpartable, categorical = lavoptions$.categorical,  :
  lavaan WARNING: parameter table does not contain thresholds
3: In lav_partable_check(lavpartable, categorical = lavoptions$.categorical,  :
  lavaan WARNING: parameter table does not contain thresholds
> summary(fit4e,fit.measures=TRUE)
lavaan 0.6-12 ended normally after 61 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          15

  Number of observations                             336

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               122.266     103.304
  Degrees of freedom                                 3           3
  P-value (Chi-square)                           0.000       0.000
  Scaling correction factor                                  1.189
```

```
  Shift parameter                                   0.477
    simple second-order correction

Model Test Baseline Model:

  Test statistic                       159.311     129.914
  Degrees of freedom                         6           6
  P-value                                0.000       0.000
  Scaling correction factor                          1.237

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                0.222       0.191
  Tucker-Lewis Index (TLI)                  -0.556      -0.619

  Robust Comparative Fit Index (CFI)                       NA
  Robust Tucker-Lewis Index (TLI)                          NA

Root Mean Square Error of Approximation:

  RMSEA                                  0.344       0.316
  90 Percent confidence interval - lower         0.294       0.265
  90 Percent confidence interval - upper         0.398       0.370
  P-value RMSEA <= 0.05                      0.000       0.000

  Robust RMSEA                                        NA
  90 Percent confidence interval - lower                   NA
  90 Percent confidence interval - upper                   NA

Standardized Root Mean Square Residual:

  SRMR                                 0.199       0.199

Parameter Estimates:

  Standard errors                       Robust.sem
  Information                           Expected
  Information saturated (h1) model        Unstructured

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  ER_Neg ~
    MIL        (a)     -4.663    1.983   -2.352    0.019
  ATAhi ~
    ER_Neg     (b)      0.030    0.008    3.569    0.000
  SelfEff ~
```

```
   MIL       (e)      2.111    1.767    1.195    0.232
 ATAhi ~
   SelfEff    (f)     -0.053    0.011   -4.631    0.000
   MIL        (c)     -0.506    0.357   -1.416    0.157
 Conscientiousness ~
   MIL        (g)      2.014    1.009    1.997    0.046
 ATAhi ~
   Cnscntsnss (h)     -0.055    0.020   -2.695    0.007

Intercepts:
               Estimate  Std.Err  z-value  P(>|z|)
  .ER_Neg          58.514    2.461   23.779    0.000
  .ATAhi           12.116    1.049   11.549    0.000
  .SelfEff         35.340    2.104   16.801    0.000
  .Conscientisnss   29.639    1.219   24.316    0.000

Variances:
               Estimate  Std.Err  z-value  P(>|z|)
  .ER_Neg         196.788   15.575   12.635    0.000
  .ATAhi            5.005    0.385   13.000    0.000
  .SelfEff        125.639    9.280   13.538    0.000
  .Conscientisnss   44.442    3.417   13.008    0.000

Defined Parameters:
               Estimate  Std.Err  z-value  P(>|z|)
   ab            -0.141    0.071   -1.980    0.048
   ef            -0.111    0.098   -1.132    0.258
   gh            -0.110    0.074   -1.483    0.138
   total         -0.868    0.327   -2.653    0.008

> parameterEstimates(fit4e,standardized=TRUE)
               lhs op               rhs label    est     se      z pvalue
1            ER_Neg  ~               MIL     a  -4.663  1.983 -2.352  0.019
2             ATAhi  ~            ER_Neg     b   0.030  0.008  3.569  0.000
3           SelfEff  ~               MIL     e   2.111  1.767  1.195  0.232
4             ATAhi  ~           SelfEff     f  -0.053  0.011 -4.631  0.000
5             ATAhi  ~               MIL     c  -0.506  0.357 -1.416  0.157
6  Conscientiousness  ~               MIL     g   2.014  1.009  1.997  0.046
7             ATAhi  ~  Conscientiousness     h  -0.055  0.020 -2.695  0.007
8            ER_Neg ~~            ER_Neg       196.788 15.575 12.635  0.000
9             ATAhi ~~             ATAhi         5.005  0.385 13.000  0.000
10          SelfEff ~~           SelfEff       125.639  9.280 13.538  0.000
11 Conscientiousness ~~   Conscientiousness       44.442  3.417 13.008  0.000
12              MIL ~~               MIL         0.123  0.000     NA     NA
13           ER_Neg ~1                           58.514  2.461 23.779  0.000
14            ATAhi ~1                           12.116  1.049 11.549  0.000
```

```
15          SelfEff ~1                          35.340  2.104 16.801  0.000
16 Conscientiousness ~1                         29.639  1.219 24.316  0.000
17              MIL ~1                           1.143  0.000     NA     NA
18               ab :=               a*b    ab  -0.141  0.071 -1.980  0.048
19               ef :=               e*f    ef  -0.111  0.098 -1.132  0.258
20               gh :=               g*h    gh  -0.110  0.074 -1.483  0.138
21            total := c+(a*b)+(e*f)+(g*h) total  -0.868  0.327 -2.653  0.008
   ci.lower ci.upper  std.lv std.all std.nox
1    -8.550   -0.777  -4.663  -0.116  -0.330
2     0.014    0.047   0.030   0.178   0.178
3    -1.352    5.574   2.111   0.066   0.188
4    -0.075   -0.030  -0.053  -0.247  -0.247
5    -1.206    0.194  -0.506  -0.074  -0.211
6     0.037    3.991   2.014   0.105   0.300
7    -0.094   -0.015  -0.055  -0.153  -0.153
8   166.261  227.314 196.788   0.987   0.987
9     4.250    5.759   5.005   0.869   0.869
10  107.450  143.828 125.639   0.996   0.996
11   37.745   51.138  44.442   0.989   0.989
12    0.123    0.123   0.123   1.000   0.123
13   53.691   63.337  58.514   4.143   4.143
14   10.060   14.172  12.116   5.049   5.049
15   31.217   39.463  35.340   3.146   3.146
16   27.250   32.028  29.639   4.421   4.421
17    1.143    1.143   1.143   3.261   1.143
18   -0.281   -0.001  -0.141  -0.021  -0.059
19   -0.304    0.081  -0.111  -0.016  -0.046
20   -0.256    0.035  -0.110  -0.016  -0.046
21   -1.509   -0.227  -0.868  -0.127  -0.362
```

Appendix 9.

R Code for Bootstrapping and Multiple Regression

```
library(readr)
library(boot)
setwd("~/Library/Mobile
Documents/com~apple~CloudDocs/Documents/R_home/bookdown_org")
data1 <- read_delim("Data Stop 0_1 MIL8_11 boot.csv", ",",
                    escape_double = FALSE, trim_ws = TRUE)
# ER_Neg vs MIL with confounds
# function to obtain correlation coefficient from the data
corcoef <- function(formula, data, indices) {
  d <- data[indices,] # allows boot to select sample
  fit <- lm(formula, data=d)
  return(summary(fit)$coefficients[2,1])
}
# bootstrapping with 1000 replications
results <- boot(data=data1, statistic=corcoef,
                R=1000, formula=ER_Neg ~ MIL+ Mother_ed+
                                Father_ed+
                                Gender+
                                Age+
```

```
                                Mother_ethnicity+
                                Father_ethnicity+
                                Grade)
        # view results
        results
        plot(results)
        # get 95% confidence interval
boot.ci(results, type="bca")
```

Appendix 10.

Perceived Family Learning Attitude

The Perceived Family Learning Attitude Scale (included below) was created for this study. The scale consists of 5 questions, with two regarding parents' learning habits and the rest regarding the family as a whole. All the questions are rated on a 5-point Likert scale from 1 – Not well at all, to 5 – Extremely well. Questions 3 and 5 are reverse-scored items.

Perceived Family Learning Attitude Scale

Question 1: Please rate how well this sentence describes your father: My father is a continuous learner. He continues to learn new things through taking courses, reading books, or researching on his own.

Question 2: Please rate how well this sentence describes your mother: My mother is a continuous learner. She continues to learn new things through taking courses, reading books, or researching on her own.

Question 3: Please rate how well this sentence describes your family: My family believes that learning is something that people only do in school.

Question 4: Please rate how well this sentence describes your family: Everyone in the family is encouraged to learn something on their spare time.

Question 5: On a scale of 1 to 5, most of my family members believe that learning is...

1 – very enjoyable; 2 – enjoyable; 3 – neutral; 4 – unpleasant; 5 – very unpleasant.

References

Associate of the Royal Schools of Music. (n.d.). Retrieved January 2, 2023, from https://us.abrsm.org/en/our-exams/diplomas/music-performance/arsm-associate-of-the-royal-schools-of-music/

Altanopoulou, P., & Tselios, N. (2018). Big Five Personality Traits and academic learning in wiki-mediated collaborative activities: Evidence from four case studies. *International Journal of Distance Education Technologies*, *16*(3), 81–92. https://doi.org/10.4018/IJDET.2018070105

Ampel, B. C., Muraven, M., & McNay, E. C. (2018). Mental work requires physical energy: Self-control is neither exception nor exceptional. *Frontiers in Psychology*, *9*, 1005. https://doi.org/10.3389/fpsyg.2018.01005

Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, *75*(3), 729–750. https://doi.org/10.1037/0022-3514.75.3.729

Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., & Martin-Skurski, M. E. (2006). Neurobiological substrates of dread. *Science*, *312*(5774), 754–758. https://doi.org/10.1126/science.1123721

The Big Five Inventory. (2009). Berkeley Personality Lab. https://www.ocf.berkeley.edu/%7Ejohnlab/bfi.htm

Blunt, A. K., & Pychyl, T. A. (2000). Task aversiveness and procrastination: A multi-dimensional approach to task aversiveness across stages of personal projects. *Personality and Individual Differences*, *28*(1), 153–167. https://doi.org/10.1016/S0191-8869(99)00091-4

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, *17*(3), 303–316. https://doi.org/10.1177/0049124189017003004

Briganti, G., Scutari, M., & McNally, R. J. (2022). A tutorial on bayesian networks for psychopathology researchers. *Psychological Methods*. https://doi.org/10.1037/met0000479

Bugos, J. A., & Cooper, P. (2019). The effects of mallet training on self-efficacy and processing speed in beginning adult musicians. *Research Perspectives in Music Education*, *20*(1), 21–32.

Canty, A., & Ripley, B. D. (2022). *boot: Bootstrap R (S-Plus) Functions*.

Certificate of Merit®. (n.d.). *Music Teachers' Association of California®*. Retrieved January 2, 2023, from https://www.mtac.org/programs/performance/certificate-of-merit/

Charalambous, A., Giannakopoulou, M., Bozas, E., & Paikousis, L. (2019). Parallel and serial mediation analysis between pain, anxiety, depression, fatigue and nausea, vomiting and retching within a randomised controlled trial in patients with breast and prostate cancer. *BMJ Open*, *9*(1), e026809. https://doi.org/10.1136/bmjopen-2018-026809

Cheval, B., Tipura, E., Burra, N., Frossard, J., Chanal, J., Orsholits, D., Radel, R., & Boisgontier, M. P. (2018). Avoiding sedentary behaviors requires more cortical resources than avoiding physical activity: An EEG study. *Neuropsychologia*, *119*, 68–80. https://doi.org/10.1016/j.neuropsychologia.2018.07.029

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Corrigall, K. A., & Schellenberg, E. G. (2015). Predicting who takes music lessons: Parent and child characteristics. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00282

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00222

Costa, P. T., & McCrae, R. R. (1992). NEO PI-R. The revised NEO personality inventory. *Psychological Assessment Resources*.

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music*, *32*(2), 139–152. https://doi.org/10.1177/0305735604041491

Covay, E., & Carbonaro, W. (2010). After the bell: Participation in extracurricular activities, classroom behavior, and academic achievement. *Sociology of Education*, *83*(1), 20–45. https://doi.org/10.1177/0038040709356565

Croyle, R. T., & Cooper, J. (1983). Dissonance arousal: Physiological evidence. *Journal of Personality and Social Psychology*, *45*(4), 782–791. https://doi.org/10.1037/0022-3514.45.4.782

Daly, H. R., & Hall, M. D. (2018). Not all musicians are created equal: Statistical concerns regarding the categorization of participants. *Psychomusicology: Music, Mind, and Brain*, *28*(2), 117–126. https://doi.org/10.1037/pmu0000213

de Roos, C., Greenwald, R., den Hollander-Gijsman, M., Noorthoorn, E., van Buuren, S., & de Jongh, A. (2011). A randomised comparison of cognitive behavioural therapy (CBT) and eye movement desensitisation and reprocessing (EMDR) in disaster-exposed children. *European Journal of Psychotraumatology*, *2*(1), 5694-Article 5694. https://doi.org/10.3402/ejpt.v2i0.5694

Dang, A. T. (2021, May 11). *Exploratory Factor Analysis in R*. Medium. https://towardsdatascience.com/exploratory-factor-analysis-in-r-e31b0015f224

DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, *11*(3), 189–228. https://doi.org/10.1214/ss/1032280214

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, *67*(3), 382–394. https://doi.org/10.1037/0022-3514.67.3.382

Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford University Press.

Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology*, *46*(6), 1454–1480. https://doi.org/10.1037/a0020204

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *The American Psychologist*, *56*(3), 218–226.

Galloway, M., Conner, J., & Pope, D. (2013). Nonacademic effects of homework in privileged, high-performing high schools. *The Journal of Experimental Education*, *81*(4), 490–510. https://doi.org/10.1080/00220973.2012.745469

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240–9245. https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003

Goroshit, M., & Hen, M. (2021). Academic procrastination and academic performance: Do learning disabilities matter? *Current Psychology (New Brunswick, N.J.)*, *40*(5), 2490–2498. https://doi.org/10.1007/s12144-019-00183-3

Guhn, M., Emerson, S. D., & Gouzouasis, P. (2020). A population-level analysis of associations between school music participation and academic achievement. *Journal of Educational Psychology*, *112*(2), 308–328. https://doi.org/10.1037/edu0000376

Guzzetti, B. J. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, *28*(2), 116–159. https://doi.org/10.2307/747886

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological Science*, 19(2), 103–108. https://doi.org/10.1111/j.1467-9280.2008.02053.x

Hajian, S. (2019). Transfer of learning and teaching: A review of transfer theories and effective instructional practices. *IAFOR Journal of Education*, *7*(1), 93–111. https://doi.org/10.22492/ije.7.1.06

Hille, A., & Schupp, J. (2015). How learning a musical instrument affects the development of skills. *Economics of Education Review*, *44*, 56–82. https://doi.org/10.1016/j.econedurev.2014.10.007

Holochwost, S. J., Bose, J. H., Stuk, E., Brown, E. D., Anderson, K. E., & Wolf, D. P. (2021). Planting the seeds: Orchestral music education as a context for fostering growth mindsets. *Frontiers in Psychology*, *11*, 586749–586749. https://doi.org/10.3389/fpsyg.2020.586749

Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of Experimental Child Psychology*, *144*, 84–97. https://doi.org/10.1016/j.jecp.2015.11.009

Jaschke, A. C., Honing, H., & Scherder, E. J. A. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Frontiers in Neuroscience*, *12*, 103–103. https://doi.org/10.3389/fnins.2018.00103

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory—Versions 4a and 54. *Berkeley, CA: University of California, Berkeley, Institue of Personality and Social Research*.

John, O. P., Robins, R. W., & Pervin, L. A. (2008). *Handbook of personality: Theory and research* (3rd ed.). Guilford Press.

Jourard, S. M., & Remy, R. M. (1955). Perceived parental attitudes, the self, and security. *Journal of Consulting Psychology*, *19*, 364–366. https://doi.org/10.1037/h0049318

Juniu, S., Tedrick, T., & Boyd, R. (1996). Leisure or work?: Amateur and professional musicians' perception of rehearsal and performance. *Journal of Leisure Research*, *28*(1), 44–56. https://doi.org/10.1080/00222216.1996.11949760

Kashahu, Dibra, Osmanaga, & Bushati. (2014). The relationship between parental demographics, parenting styles and student academic achievement. *European Scientific Journal (Kocani)*, *10*(13).

Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology*, *93*(1), 59–69. https://doi.org/10.1037/0021-9010.93.1.59

Keith, N., Horvath, D., & Klamar, A. (2020). The more severe the merrier: Severity of error consequences stimulates learning from error. *Journal of Occupational and Organizational Psychology*, *93*(3), 712–737. https://doi.org/10.1111/joop.12312

King, R. B., & McInerney, D. M. (2014). The work avoidance goal construct: Examining its structure, antecedents, and consequences. *Contemporary Educational Psychology*, *39*(1), 42–58. https://doi.org/10.1016/j.cedpsych.2013.12.002

Kool, W., & Botvinick, M. (2013). The intrinsic cost of cognitive control. *The Behavioral and Brain Sciences*, *36*(6), 697-. https://doi.org/10.1017/SO140525X1300109X

Kralovec, E., & Buell, J. (2000). *The end of homework: How homework disrupts families, overburdens children, and limits learning*. Beacon Press.

Kruse-Weber, S., & Parncutt, R. (2014). Error management for musicians: An interdisciplinary conceptual framework. *Frontiers in Psychology*, *5*, 777–777. https://doi.org/10.3389/fpsyg.2014.00777

Kuckelkorn, K. L., de Manzano, Ö., & Ullén, F. (2021). Musical expertise and personality – differences related to occupational choice and instrument categories. *Personality and Individual Differences*, *173*, 110573. https://doi.org/10.1016/j.paid.2020.110573

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70. https://doi.org/10.1016/j.copsyc.2015.08.003

Lane, A. M., Jones, L., & Stevens, M. J. (2002). Coping with failure: The effects of self-esteem and coping on changes in self efficacy. *Journal of Sport Behavior*, *25*(4), 331–345.

Lane, J., Lane, A. M., & Kyprianou, A. (2004). Self-efficacy, self-esteem and their impact on academic performance. *Social Behavior and Personality*, *32*(3), 247–256. https://doi.org/10.2224/sbp.2004.32.3.247

Lay, C. (1992). Trait procrastination and the perception of person-task characteristics. *Journal of Social Behavior and Personality*, *7*(3), 483–494.

Lay, C. H., & Brokenshire, R. (1997). Conscientiousness, procrastination, and person-task characteristics in job searching by unemployed adults. *Current Psychology*, *16*(1), 83–96. https://doi.org/10.1007/s12144-997-1017-9

Leutgeb, V., Schäfer, A., & Schienle, A. (2009). An event-related potential study on exposure therapy for patients suffering from spider phobia. *Biological Psychology*, *82*(3), 293–300. https://doi.org/10.1016/j.biopsycho.2009.09.003

Lonergan, J. M., & Maher, K. J. (2000). The relationship between job characteristics and workplace procrastination as moderated by locus of control. *Journal of Social Behavior and Personality*, *15*(5), 213–224.

Loui, P., Raine, L. B., Chaddock-Heyman, L., Kramer, A. F., & Hillman, C. H. (2019). Musical instrument practice predicts white matter microstructure and cognitive abilities in childhood. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01198

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, *1*(2), 75–86. https://doi.org/10.1093/scan/nsl013

Margaritis, D. (2003). *Learning Bayesian Network Model Structure from Data*. (Publication No. CMU-CS-03-153) [Ph.D dissertation, Carnegie Mellon University]. Carnegie Mellon University School of Computer Science. https://www.cs.cmu.edu/~dmarg/Papers/PhD-Thesis-Margaritis.pdf

Marjoribanks, K. (1987). Ability and attitude correlates of academic achievement: Family-group differences. *Journal of Educational Psychology*, *79*(2), 171–178. https://doi.org/10.1037/0022-0663.79.2.171

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, *86*, 95–104. https://doi.org/10.1016/j.brat.2016.06.006

McNally, R. J., Robinaugh, D. J., Deckersbach, T., Sylvia, L. G., & Nierenberg, A. A. (2022). Estimating the symptom structure of bipolar disorder via network analysis: Energy dysregulation as a central symptom. *Journal of Psychopathology and Clinical Science*, *131*(1), 86–97. https://doi.org/10.1037/abn0000715

Milgram, N. A., Sroloff, B., & Rosenbaum, M. (1988). The procrastination of everyday life. *Journal of Research in Personality*, *22*(Jun 88), 197–212.

Milgram, N., Marshevsky, S., & Sadeh, C. (1995). Correlates of academic procrastination: Discomfort, task aversiveness, and task capability. *The Journal of Psychology*, *129*(2), 145–155. https://doi.org/10.1080/00223980.1995.9914954

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., & Lee, Y.-H. (2011). Mind your errors: Evidence for a neural mechanism linking growth mind-set to adaptive post error adjustments. *Psychological Science*, *22*(12), 1484–1489. https://doi.org/10.1177/0956797611419520

Müllensiefen, D., & Harrison, P. (2020). The impact of music on adolescents' cognitive and socio-emotional learning. In *the "braincando" handbook of teaching and learning*. David Fulton Publishers.

Onwuegbuzie, A. J., & Collins, K. M. T. (2001). Writing apprehension and academic procrastination among graduate students. *Perceptual and Motor Skills*, *92*(2), 560–562. https://doi.org/10.2466/pms.2001.92.2.560

Palmer, J. C., & Strickland, J. (2016, June). *A beginner's guide to crowdsourcing*. Https://Www.Apa.Org. https://www.apa.org/science/about/psa/2016/06/changing-minds

Pearl, J. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons Ltd.

Perera, L. D. H. (2014). Parents’ attitudes towards science and their children’s science achievement. *International Journal of Science Education*, *36*(18), 3021–3041. https://doi.org/10.1080/09500693.2014.949900

Pope, D. (2010). Beyond ‘doing school’: From ‘stressed-out’ to ‘engaged in learning.’ *Education Canada*, *50*(1), 4–8.

Regression with categorical variables: Dummy coding essentials in R - Articles - STHDA. (n.d.). Retrieved January 21, 2023, from http://www.sthda.com/english/articles/40-regression-analysis/163-regression-with-categorical-variables-dummy-coding-essentials-in-r/

Revelle, W. (2022). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roden, I., Könen, T., Bongard, S., Frankenberg, E., Friedrich, E. K., & Kreutz, G. (2014). Effects of music training on attention, processing speed and cognitive music abilities-findings from a longitudinal study. *Applied Cognitive Psychology*, *28*(4), 545–557. https://doi.org/10.1002/acp.3034

Rosseel, Y. (2012). **lavaan**: An *R* package for structural equation modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, *20*, 55–67. https://doi.org/10.1016/j.edurev.2016.11.005

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23–74.

Schniederjans, M. J., & Kim, E. B. (2005). Relationship of student undergraduate achievement and personality characteristics in a total web-based environment: an empirical study. *Decision Sciences Journal of Innovative Education*, *3*(2), 205–221. https://doi.org/10.1111/j.1540-4609.2005.00067.x

Scutari, M. (2010). *Learning Bayesian networks with the bnlearn R Package* (arXiv:0908.3817). arXiv. https://doi.org/10.48550/arXiv.0908.3817

Senecal, C., Lavoie, K., & Koestner, R. (1997). Trait and situational factors in procrastination: An interactional model. *Journal of Social Behavior and Personality*, *12*(4), 889–903.

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, *81*(395), 826–831. https://doi.org/10.1080/01621459.1986.10478341

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, *51*(2), 663–671. https://doi.org/10.2466/pr0.1982.51.2.663

Silas, S., Müllensiefen, D., Gelding, R., Frieler, K., & Harrison, P. M. C. (2022). The associations between music training, musical working memory, and visuospatial working memory. *Music Perception*, *39*(4), 401–420. https://doi.org/10.1525/mp.2022.39.4.401

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, *87*, 287. https://doi.org/10.1111/j.2044-8295.1996.tb02591.x

Solomon, L. J., & Rothblum, E. D. (1984). Academic procrastination: Frequency and cognitive-behavioral correlates. *Journal of Counseling Psychology*, *31*(4), 503–509. https://doi.org/10.1037/0022-0167.31.4.503

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: structure, development, and outcomes. *Current Directions in Psychological Science*, *24*(5), 358–362. https://doi.org/10.1177/0963721415589345

Steel, P. (2007). The Nature of Procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, *133*(1), 65–94. https://doi.org/10.1037/0033-2909.133.1.65

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thalheimer, W., & Cook, S. (2009). *How to calculate effect sizes from published research: A simplified methodology*.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

Trautwein, U., Niggli, A., Schnyder, I., & Lüdtke, O. (2009). Between-teacher differences in homework assignments and the development of students' homework effort, homework emotions, and achievement. *Journal of Educational Psychology*, *101*(1), 176–189. https://doi.org/10.1037/0022-0663.101.1.176

Vatterott, C. (2018). *Rethinking homework: Best practices that support diverse needs*. Association for Supervision & Curriculum Development. http://ebookcentral.proquest.com/lib/harvard-ebooks/detail.action?docID=5704245

Zhao, B. (2011). Learning from errors: The role of context, emotion, and personality. *Journal of Organizational Behavior*, *32*(3), 435–463. https://doi.org/10.1002/job.696