

Affective Neuroscience meets Labor Economics: Assessing Non-Cognitive skills on Late Stage Investment on at-Risk Youth*

Pablo A. Egana del Sol[†]

Columbia University

August, 2016

*For comments and discussion I thank Miguel Urquiola, Elke Weber, Mark Dean, Paul Sajda, Cristian Pop-Eleches, Pietro Ortoleva, Michael Woodford, Joseph Stiglitz, Andres Velasco, Chris Blattman, Ricardo Hausmann, Cesar Hidalgo, Karla Hoff, Sanjay Peters, Douglas Almond, Eric Verhoogen, Ernesto Reuben, Alessandra Casella, John Mutter, Mel Win Khaw, Scott Weiner, Paulo Barraza, Michael Platt, Esteban Puentes, Dante Contreras, Francisco Gallego, Paula Luengo, Hugo Nopo, Claudia Martinez and seminar participants at Columbia University, Center for International Development at Harvard University, MIT Media Lab, Michael Platt's Lab at Wharton and Perelman School of Medicine at University of Pennsylvania, InterAmerican Development Bank, D.C., University of Chile, 2014 Neuroeconomic Summer Workshop in Erice, Italy, the Cognition and Decision (Neuro-economic) Lab, and the Laboratory for Intelligent Imaging and Neural Computing (LIINC) both at Columbia University. For implementation I thank Benjamin Faivovich, Jose Miguel Barros and Esteban Alvarez from Emprende Joven in Chile. I acknowledge the financial support provided by the Centre for Social Conflict and Cohesion Studies (CONICYT/FONDAP/15130009) at University of Chile and the Inter-American Development Bank. Nicole Moscovitz and Tao Tu provided invaluable research assistance on the neuroscience paradigm. Simon Ballesteros, Macarena Gonzalez, Nicolas Fernandez and Jose Manuel Belmar provided research assistance in the field.

[†]Email: pae2110@columbia.edu.

Abstract

There is a common puzzle in the labor economic literature in which social programs —i.e. educational or labor market programs— that exhibit positive impacts on labor market outcomes seem, contrary to expectations, not to affect measures of non-cognitive skills (Kautz et al., 2014; Calero et al., 2014; West et al., 2015). Furthermore, recent meta-analysis concludes that studies using self-reported psychometric tests to proxy socio-emotional skills face potentially significant measurement errors that could lead to inconsistent impact evaluations (Almlund et al., 2011; Heckman and Kautz, 2014). Hence, this paper seeks to contribute in three dimensions. First, using novel data, it provides a rigorous estimate of the impact of a program to foster creative and life skills for at-risk youth in a formal educational setting in a developing country context. Second, using neurophysiological and survey data from field experiments, it assesses the impact of that program on emotional state. Third, using neurophysiological recordings, it measures the program’s impact on emotional responsiveness as a proxy for behavioral response (Loewenstein, 2001; DellaVigna, 2009). There are three main findings. First, I find that the program had significant impact on educational outcomes —i.e. in both dropouts and SAT-like registration— yet no impact on the expected mechanism, such as socio-emotional skills or creativity measures, which is consistent with the findings in the labor economics literature referenced above. Second, I find significant impacts on emotional state —in both arousal and valence indices— from neurophysiological recordings. If emotional disposition can bias self-reported measures of non-cognitive skills, as suggested above (Querengasser and Schindler, 2014; Egana-delSol, 2016), this may account for the lack of evidence of an impact on non-cognitive skills —the puzzle highlighted above. Third, program participation also reduces individuals’ emotional reaction to negative stimuli —one might say that it makes individuals more resilient.

JEL codes: I2, I25, J24, D87

Keywords: Non-cognitive skills, randomized controlled trial, impact evaluation, education, training, emotions, EEG, emotion-detection, affective neuroscience, applied neuroscience.

Contents

- 1 Introduction** **4**

- 2 Background and Conceptual Framework** **8**
 - 2.1 Mining’s Rockstars Program 8
 - 2.2 Brief Review of Literature on Late Stage Interventions 11
 - 2.3 Education Production Function 13
 - 2.4 Definition and Measurement of Skills 15
 - 2.4.1 Skill Definitions 15
 - 2.4.2 Measurements of Non-cognitive Skills 17

- 3 The Role of Emotions** **19**
 - 3.1 Emotion-Detection Theory 22

- 4 Experimental Paradigm** **27**

- 5 Results** **30**
 - 5.1 Descriptive Statistics 30
 - 5.2 Attrition 32
 - 5.3 Impact on Educational Outcomes 34
 - 5.4 Impact on Self-reported measures of Non-cognitive Skills 36
 - 5.5 Impact on Emotional Regulation 39
 - 5.6 Discussion 41
 - 5.6.1 Resting-state Emotional State 43
 - 5.6.2 Emotional Responsiveness 46

- 6 Conclusions** **48**

- 7 Appendix 1: Balance of Attrition** **64**

- 8 Appendix 2: Robustness Checks** **64**
 - 8.1 Robustness Check: Impact on Cognitive Skills (Raven-like Test) 64
 - 8.2 Robustness Check: In Difference Model: Impact on Emotional Regulation 65
 - 8.3 Robustness Check: Wild Bootstrap on p-values (Exact p-values) 66

- 9 Appendix 3: Field Experiment Setting** **70**

1 Introduction

In developing countries, youth often face scarce employment opportunities. The literature generally cites a lack of appropriate skills as a key reason for this (Bassi and Urzua, 2010; Pierre et al., 2014). Since the work of Bowles and Gintis (1976); Gintis (1971), research holds that these skills can be both cognitive and non-cognitive in nature. Cognitive skills encompass the use of language, memory, and logical-mathematical knowledge, while non-cognitive skills —also referred to as socio-emotional or life skills— include self-confidence, self-control, internal locus control, and grit (perseverance).

Recent evidence shows that interventions aimed at the non-cognitive skills of youth can have a significant impact on behavior and/or outcomes. For instance, Carrel and Sacerdote (2013) study a college coaching intervention in New Hampshire and find a positive effect on enrollment and continued attendance in college. Blattman et al. (2015), shows that cognitive behavioral therapy can reduce crime and violence among at-risk youth in Liberia. In addition, studies suggests that programs to improve socio-emotional skills are more effective among students who are still enrolled in secondary schools (Heckman and Kautz, 2012; Cunha et al., 2010).

The existence of such impacts is not surprising to the extent that the neuroscience literature suggests that it is possible to affect non-cognitive skills during adolescence.¹ The prefrontal cortex, which is related to emotion and self-control, is malleable into the early 20s (Fuster, 2002, 2013; Sigman et al., 2014) because the brain is still developing, particularly those areas related to identity, moral and social consequences of actions, and emotions (Fuster, 2013). Indeed, during adolescence there is more synaptic activity of new neurons than in the first years of life (Koelsch, 2012; Fuster, 2002, 2013; Levitin, 2006).² To summarize, both the economics and the neuroscience literature suggest that late-stage non-cognitive investments can have a positive impact on the development of non-cognitive skills.

In this context a puzzle arises in the labor economics literature. Specifically, the evidence indicates that educational or labor market programs can improve individuals'

¹ Respect to cognitive skills is still an open debate. On one hand, some authors argue that after age 10, IQ level —one of the main measures of cognitive skills— remains relatively stable (Almlund et al., 2011; Koelsch, 2012). On the other hand, there is also evidence that shows that even cognitive skills can be affected during adolescence or later (Sigman et al., 2014).

² The brain's synapses are programmed to grow for a number of years until they shift to pruning away unneeded connections. Myelination—a substance that coats the axons, speeding up the synaptic transmission—is boosted between the ages of 8 and 16 (Koelsch, 2012).

labor market outcomes; yet, contrary to the expectation the above would suggest, they do not seem to affect individuals' non-cognitive skills (Calero et al., 2014; Card et al., 2011; Ibarra et al., 2014; West et al., 2015). This is despite the fact that the qualitative evidence arising from participant, employer, expert, and program manager interviews usually suggests that improvements in non-cognitive skills are a main channel for the labor market effects (Calero et al., 2014; Ibarra et al., 2014; Fazio, 2011). This emerges, for instance, in work on active labor market policies (ALMP) (Calero et al., 2014; Card et al., 2011; Ibarra et al., 2014) and in the literature covering in-school programs (Morrison and Shoon, 2013; Egana-delSol, 2016a; West et al., 2015).

In short, the theory and the qualitative evidence suggest that interventions' impact on labor market outcomes should be mediated by improvements in non-cognitive skill, yet there is little evidence that they actually affect socio-emotional skills. There are two main candidate explanations for this. First, it may be that non-cognitive skills simply do not account for the labor market impacts that are found. Second, it may be that the impacts on socio-emotional skills are difficult to detect because they are measured with noise, in part because they are generally self-reported. Consistent with the latter possibility, recent meta-analyses conclude that studies using self-reported psychometric tests to proxy socio-emotional skills face potentially significant measurement error.³

This study provides further evidence consistent with the second possibility. My hypothesis is that social programs affect participants' emotional regulation⁴ which in turn makes it difficult to measure their impact on non-cognitive skills.⁵

Emotional regulation matters because neuroscientists highlight the critical role that emotions play in cognition, perception, attention, and memory (Damasio, 1994; Lakoff, 2008; Salzman and Fusi, 2010; Fuster, 2013). Furthermore, work that combines neuroscience, behavioral science, and economics supports the notion that emotions influence economic behavior and decision-making, as well as labor market performance

³ See for a review Almlund et al. (2011); Heckman and Kautz (2012).

⁴ Emotional regulation is a different concept than emotional intelligence. On one hand, emotional regulation can be define as a mixture of cognitive and emotional processes that shape a mental state —i.e. a disposition to act (Salzman and Fusi, 2010). In other words, emotional regulation involves developing the ability to affect the emotional response to a stimulus. On the other hand, emotional intelligence consists of four capabilities or competencies: self awareness, self management, social awareness and social skills (Goleman, 2010).

⁵ In Egana-delSol (2016b) I study the relation between transient emotional state and self-reporting on non-cognitive skills tests. See also Querengsser and Schindler (2014).

(e.g., occupational choice, salaries, entrepreneurship, etc.).⁶ Taken together, this suggests that emotion can influence cognition and behavior in powerful ways.⁷ Indeed, the behavioral economics literature sustains that even minor mood manipulations have a substantial impact on behavior (DellaVigna, 2009).

To explore these issues, I follow the James-Lange theory of emotions, in which emotions are organized along the arousal and valence locus. Arousal is related to excitement. Valence could be interpreted as a positive or negative mood, as well as an attitude of either approach or withdrawal towards/from a stimulus (Harmon-Jones et al., 2010; Kassam et al., 2013). To proxy emotional regulation —i.e. emotional state and responsiveness— I rely on emotion-detection theory from the affective neuroscience literature that uses electroencephalogram (EEG) recordings to measure emotions. In particular, I use low-cost portable EEG (EEG) headsets to obtain a proxy measure of subjects’ emotional states — i.e. pre-test resting state— as well as emotional responsiveness to both positive and negative stimuli, in the arousal-valence locus.⁸

In particular, I study the Chilean program “Mining’s Rockstars” (“Rockstars para la Minería” in Spanish), which aims to foster life skills with a focus on self-confidence and creativity, following principles that are summarized as: “learning by failing, gaming, doing, and rethinking”. The program is implemented in the IV Coquimbo Region of Chile, where mining is the central economic activity. “Mining’s Rockstars” creates didactic materials, namely: a student textbook, a teacher textbook, and class-by-class guidelines together with videos for each activity, all based on the aforementioned principles and a Harvard leadership model (McClelland, 1973). The intervention consists of weekly workshops in which students participate in different activities designed to improve their life skills. Program instructors work together with school teachers to train them and enhance local capabilities at the schools.

The program was randomly assigned to technical/vocational high schools enrolling students aged 16-18. I use this set-up to explore the effects that students’ participation (for one academic semester—around four months) has on educational outcomes — e.g. SAT-like test registration and dropout rates— and on their non-cognitive skills.⁹

⁶ For recent reviews see Weber and Johnson (2009); Lempert and Phelps (2014); Lerner et al. (2015).

⁷ For reviews see Loewenstein et al. (1992, 2001); Damasio (1994); Lerner et al. (2015); Weber and Johnson (2009); Lempert and Phelps (2014).

⁸ See for example Takahashi et al. (2004); Ramirez and Vamvakousis (2012); Verma and Tiwary (2014); Brown et al. (2011); Bos (2006).

⁹ I also observe the effects on cognitive skills using the Raven’s test (Raven, 1936); I use this score

I measure these using three tests: Rotter’s locus control scale, grit—perseverance—Scale, and Torrance’s test of creative thinking.¹⁰ In addition, I explore how program participation affects emotional regulation.

There are three main findings. First, I find that the program had significant impact on educational outcomes —i.e. in both dropouts and SAT-like registration— yet no impact on the expected mechanism, such as socio-emotional skills or creativity measures, which is consistent with the findings in the labor economics literature referenced above. Second, I find significant impacts on emotional state—in both arousal and valence indices— from neurophysiological recordings. If emotional disposition can bias self-reported measures of non-cognitive skills, as suggested above (Querengsser and Schindler, 2014; Egana-delSol, 2016b), this may account for the lack of evidence of an impact on non-cognitive skills—the puzzle highlighted above. Third, program participation also reduces individuals’ emotional reaction to negative stimuli—one might say that it makes individuals more resilient.

The present study involves different scientific disciplines, including the economics of education, labor economics, affective neuroscience, and applied neuroscience —i.e. brain-computer interface. Given the specificity, sensitivity and complexity of the required dimensions of this study, the data were directly collected by the author. To the best of my knowledge, this is the first research study to apply large-scale neurophysiological recordings from electroencephalograms (EEG) in field experiments, and also in the context of a social program evaluation. In short, the purpose of this study is to contribute to the understanding of the human capital production function in terms of cognitive, socio-emotional and creative skills by combining neurophysiological, behavioral and demographic data in a late-stage intervention setting.¹¹

The methodology proposed in this paper has many benefits. First, it offers a way to incorporate emotion into the labor economics field. The importance of emotional stability in labor markets and overall life satisfaction has recently been highlighted for both developed and developing countries (Deming, 2015; OECD, 2015). Furthermore,

as a placebo. See the appendix for details.

¹⁰ For details see the next section or Grit Scale (Duckworth et al., 2007), Locus Control Scale (Rotter, 1966), and Torrance’s test of creative thinking (Kim et al., 2013; Kaufman and Sternberg, 2010).

¹¹ Regardless, since I am doing a program evaluation in the medium scale range, this study will not address equilibrium effects, which are more likely to occur when the program or policy are taken to a scale and sustained for a longer period of time (Banerjee and Duflo, 2008; Acemoglu, 2010; Deaton, 2010; Pop-Eleches and Urquiola, 2013).

international multilateral institutions have emphasized the need to improve cognitive and socio-emotional skill assessment.¹²

Second, it also may aid evaluation of similar programs attempting to foster non-cognitive skills. In fact, the effects on emotional disposition and responsiveness also have implications for the experimental evaluation of educational interventions. Social programs for education and Active Labor Market Policies (ALMP) usually aim to impact socio-emotional or life-skills, such as perseverance, self-control, goal-oriented effort, and so forth. However, these factors suffer from measurement bias —i.e. reference or emotional bias— due to self-reporting. The methodology proposed here allows us to measure emotional disposition and responsiveness from EEG recordings, which is a non-invasive and low-cost method. Therefore, further research could incorporate physiological measures of emotional regulation to study the human capital production function, educational interventions, and the ALMP effectiveness.

The paper is organized as follows: section 2 defines the background and conceptual framework for education production function and skill measurement. Section 3 describes the theories of emotions, their implications on behavior, and their methods to be measured. Section 4 explains the methodology, experiment and data, and section 5 presents the results and a discussion about the main findings. Finally, conclusions are presented in section 6.

2 Background and Conceptual Framework

This section has four objectives: to explain and contextualize the intervention studied; to review the literature on late-stage interventions aiming to foster non-cognitive skills; to sketch a conceptual framework relative to the education production function; and to delineate and discuss the definitions and measurements of non-cognitive skills.

2.1 Mining’s Rockstars Program

Technical and Professional Education (TPE), which is similar to vocational-technical schools in the US, started in Chile in the late 60s as an educational alternative oriented

¹² For instance, the Organization of Economic Cooperation and Development (OECD, 2015), the World Bank (McKenzie et al., 2014; McKenzie, 2014), and the Inter American Development Bank (Ibarraran et al., 2014).

to working life, but with the possibility of continuing to higher education. It began as a four year cycle after completion of elementary school education. In the early 80s, TPE was divided into two blocks: the first two years dedicated to normal high school programming, known as Scientific and Humanistic Education (SHE), and the next two years focused on TPE.

After the 1981 education liberalization reform, schools were given the ability to choose whether they wanted to offer SHE or TPE, and to manage the content of each track (Cox, 2006). The goal of allowing high school managers —i.e. the principal and the municipality education department (DAEM) officer¹³— to choose among tracks was to improve the quality and pertinence of the education offered, though evidence shows that this was not necessarily the result (Albarran and Gonzalez, 2015).

Nowadays, 42% of students in TPE programs continue on to higher education, compared to 66% of the students in SHE programs (Bassi and Urzua, 2010). Only 39% of those TPE students complete their degree. This suggests that students of TPE are disadvantaged in terms of access to and completion of higher education, which is reflected in the achievements of these students in several dimensions. In fact, after controlling for demographics and institutional characteristics, TPE schools underperform SHE schools in grades, access to higher education, completion rates, among others (Valenzuela et al., 2013). In addition, employers, in particular those in the mining sector, argue that TPE students lack employability skills, such as self-initiative, ability to finish a task well, ability to work as a team, tendency to honor commitments, and so forth (Albarran and Gonzalez, 2015). The National Mining Council also established a Competencies Framework for the mining sector, with its first priority is the development of life skills among applicants for jobs in the industry (Albarran and Gonzalez, 2015).

The Mining’s Rockstars program was created to tackle these challenges. The program aims to foster ten different skills through a didactic weekly intervention following a recognized Harvard leadership model by McClelland (1973) with the principles of “learning by failing, gaming, doing and rethinking.”¹⁴ These skills include: networking

¹³DAEM is the Spanish acronym of Municipal Education Department (Departamento de Administración de Educación Municipal). The DAEM is the institution within municipalities in charge of provide public education in Chile after the 1981 reform.

¹⁴ McClelland (1973) tested more than 500 businesses, sports, religious and political leaders of the world with the intention to see what skills they addressed made them different from normal people. He found 30 non-cognitive skills that were relevant, where 10 are developable in the short term. These

and mutual support, information searching, calculated risk taking, accomplishment of work-related commitments, systematic planning and monitoring, persuasion, demand for high quality and efficient work, and self-confidence. The intervention has a particular focus on self-confidence and creativity.

Accordingly, the program created didactic materials to be used by students and teachers, including a student textbook, which is hardcover, full color, and has more than 200 pages, and a teacher textbook, which has class-by-class guidelines and videos for each activity. The program’s coaches work together with teachers in order to generate local capabilities. The program also includes three training seminars and pilot/simulated training workshops before the program starts, as well as an online web support platform for teachers to use during implementation.

As a part of the intervention, the classroom is transformed into a friendly game room. All school desks are put aside, and chairs are arranged in a half-moon form, so that each student sits the same distance from the center. Then, through games and reflections/deliberations, working on teams under pressure, addressing complex tasks and setting their own goals to achieve through these tasks, students are taught to recognize their own strengths and weaknesses.

The intervention evaluated here was a cluster-level —i.e. school level— randomization that targeted young students who were in technical/vocational high school for one academic semester, lasting around four months. Since there is high regional heterogeneity among schools in Chile, only TPE schools belonging to the 4th Region of Coquimbo were considered.¹⁵ The program’s managers agreed with municipal education department (DAEM) authorities on a potential subset of schools where it was feasible and desirable to implement the program. Restrictions included having more than 50 students per cohort and the guaranteed participation of the school’s principal in the program. Finally, “Mining’s Rockstars” managers randomly selected a group of four schools to be treated, as there were not sufficient resources to apply the program to all schools in the region. The annual estimated budget of the program was around USD\$100,000, and the Ministry of Education agreed to fund only a pilot.¹⁶

10 skills are: 1. Search for the opportunity and take initiative; 2. Persistence (grit); 3. Work in support networks; 5. Take calculated risk; 6. Comply work commitments; 7. Systematic planning and monitoring; 8. Persuasion; 9. Demanding efficiency and quality 10. Self-Confidence.

¹⁵ Valenzuela et al. (2013).

¹⁶ The fixed cost —methodology design, experts, professionals, book design and videos— was 64%, while the variable cost —books, workshops, monitoring, etc — account for the rest. The estimated cost of the book was USD\$ 24.

Within the treatment schools, the program was open to those students in their 4th and last year of high school. Therefore, I collected data on 4th grade students in both treatment and control group schools. As an additional exercise, I considered another control group consisted of students who attended same schools and were in their 3rd year. Since the intervention targeted only students in 4th grade, it was natural to create a within school control group as a way to perform additional robustness check on the results. Finally, the sample with valid EEG measurements is comprised of 296 students in total, of which 140 belong to the treatment group.¹⁷ The empirical strategy is explained in a following section.

2.2 Brief Review of Literature on Late Stage Interventions

This subsection reviews recent studies on late-stage interventions that are relevant to the present study.¹⁸ Recent evidence shows that late-stage non-cognitive interventions on youth can have significant impacts in terms of behavior and/or outcomes (Blattman et al., 2015; Calero et al., 2014; Carrel and Sacerdote, 2013). However, evidence concerning the returns from late-stage non-cognitive investments on academic and/or labor market outcomes is scarce.¹⁹ There are only a few rigorous evaluations of interventions of this type in the context of developing countries.²⁰

A couple of experimental studies conducted during the last few years in Latin

¹⁷ I collected almost 500 EEG recordings in students. Filtering for artifacts and data quality gives us the final sample. See Attrition subsection for details.

¹⁸ In addition, since I am doing a program evaluation in the medium scale range, this study will not address equilibrium effects, which are more likely to occur when the program or policy are taken to a scale and sustained for a longer period of time (Banerjee and Duflo, 2008; Acemoglu, 2010; Deaton, 2010; Pop-Eleches and Urquiola, 2013).

¹⁹ There is a literature focused on Active Labor Market Policies (ALMP) focused in industrialized countries, with few studies related to developing countries, in particular, in the context of Latin America. Examples of ALMPs are: demand incentives, such as targeted wage subsidies or tax incentives, public employment, flexible employment arrangement such as remote work or internship, and training and vocational programs. For examples on these ALMP see the extensive review of Puentes and Urzua (2010) for Latin American countries, or Card et al. (2010), Kluve (2010), and Card et al. (2015) for industrialized countries context. In particular, in the industrialized countries context, the main conclusion is that ALMPs have low effectiveness in term of both earnings and employment in the developed world, with classroom and on-the-job training being the most effective settings (Card et al., 2010). Nonetheless, compared to industrialized countries, developing countries, in particular those in Latin American, unveil greater effectiveness, especially among women, although the absolute magnitude is yet small (Puentes and Urzua, 2010).

²⁰ As mentioned in the introduction, there is recent evidence in a developed country context. In particular, Carrel and Sacerdote (2013) study a college coaching intervention in New Hampshire. The authors find a positive and significant impact of coaching on enrollment and continuing in college.

America are closely related to the present study. Specifically, [Card et al. \(2011\)](#) conducted one of the first rigorous studies of the effectiveness of a labor training program that fosters vocational and socio-emotional skills, in the context of the Dominican Republic’s Youth and Employment —Juventud y Empleo— program. [Attanasio et al. \(2015\)](#) studied a similar program in Colombia. Both studies show similar results: no impact on employment and only a modest impact on earnings and quality of employment, conditional on working.

[Calero et al. \(2014\)](#) study the effectiveness of an art-based program to improve employability among at-risk youth in Brazil’s slums. The program jointly employs expressive art and theater training with mainstream vocational and academic instruction, aiming to improve labor market outcomes such as earnings and obtainment of paid formal employment, among others. The authors do not find robust and statistically significant impacts on non-cognitive measures, although their own qualitative data shows that there are impacts on valuable labor market skills such as punctuality, responsibility, and dedication to work. They point out that this lack of quantitative impact could be due to selection problems as well as to the self-reported nature of the psychometric tests employed.²¹

In a follow up study of the same Youth and Employment program, [Ibarraran et al. \(2014\)](#) found similar results. Interestingly, they incorporated measurements of life skills, including personality, grit and self-esteem. Using customized instruments to measure those skills, the authors find a positive but moderate impact on some dimensions. They also remark that improvements in the instruments to measure life skills are needed. Other authors have similar conclusions ([Almlund et al., 2011](#); [Calero et al., 2014](#); [Heckman and Kautz, 2012](#)). For instance, recent meta-analysis conducted by [Heckman and Kautz \(2012\)](#) concludes that studies using self-reported psychometric tests to proxy socio-emotional skills face potentially significant measurement error with respect to the tests used. Moreover, [Cunha and Heckman \(2007\)](#) point out that measurement error is high in all of the proxies that they are using to measure cognitive and non-cognitive skills in a latent factor model.²² Indeed, the same authors argue that slight changes in those measures substantially affect the direction and significance of their findings.

²¹ The Social and Personal Competencies Scale, an adaptation of the Big Five Inventory done by [Brea \(2010\)](#); and the Grit Scale, developed by [Duckworth et al. \(2007\)](#).

²² See [Egana-delSol \(2016b\)](#) for details on this type of model.

The most recent study is conducted by [Blattman et al. \(2015\)](#), who evaluates a late-stage non-cognitive intervention in at-risk youth in Liberia. In his assessment of a youth population recruited in the streets, the authors find significant results on reducing violence, crime and poverty indicators from a cognitive behavioral therapy intervention. The main conclusion is that orthodox policies to increase labor participation of at-risk youth by increasing police control and low-skill emergency job creation had failed, and that there is a room for late investment in non-cognitive skills in order to change youth behavior.

2.3 Education Production Function

I used a production function to describe the process through which the skills of students evolve between the beginning and the end of the program. The baseline period is indexed as 1, at the beginning of the academic year, in March of 2014, while the post-intervention period is identified as 2, when students finish one semester of the program, in June of 2014. Students' outcomes or skills at time 2 are assumed to be a function of previous investments and some other shocks.

Moreover, relevant inputs in the production and investment functions will depend on the intervention. In this sense, the model allows for changes in both the parameters—i.e. elasticities or relative prices—and parent/student behavior due to participation in the program.²³

The main contrastable groups are the following: the treated group comprises of those students who participated in the program for one semester; and the control, or non-treated, group includes students who did not participate in the program, and thus were enrolled in neighboring schools.

This type of model allows the exploration of general equilibrium effects, or behavioral responses from students, principals and families. I did not collect information about parents' investment or school adjustments, therefore, the estimation of the model will contemplate only individual investments and behaviors. The conceptual framework will follow [Todd and Wolpin \(2003\)](#) and [Pop-Eleches and Urquiola \(2013\)](#), where achievement depends on the school and previous family investment. This conceptual framework will be later adapted to the context of this study. In fact, I consider that

²³ For details see [Todd and Wolpin \(2003\)](#).

the student's achievement —e.g. on non-cognitive tests— at the beginning of time 1 is:

$$A_1 = g_0(F_0, \mu), \quad (1)$$

where F and μ index the her family's investments and her innate ability, respectively. Then, in the second period we have:

$$A_2 = g_1(S_1, F_1, F_0, \mu) \quad (2)$$

where S accounts for school's inputs in the previous period.

Now, define the parent's expected school's inputs as:

$$\bar{S}_1 = \theta(A_1, W, \mu) \quad (3)$$

where W denotes family wealth. The actual school's input at period 1 is defined as:

$$S_1 = \psi(A_1, \mu) \quad (4)$$

Finally, family's input is defined by:

$$F_1 = \phi(A_1, W, \mu, S_1 - \bar{S}_1) \quad (5)$$

where the family observed $(S_1 - \bar{S}_1)$ before set the household investment F_1 .

The Production Function Effect (i.e. *ceteris paribus*) of a change in S_1 will be:

$$\frac{\partial A_2}{\partial(S_1 - \bar{S}_1)} = \frac{\partial A_2}{\partial S_1} = \frac{\partial g_1}{\partial S_1} \quad (6)$$

While the aggregate Policy Effect (with behavioral adjustment):

$$\frac{\partial A_2}{\partial(S_1 - \bar{S}_1)} = \frac{\partial A_2}{\partial S_1} = \frac{\partial g_1}{\partial S_1} + \frac{\partial g_1}{\partial F_1} \frac{\partial F_1}{\partial(S_1 - \bar{S}_1)} \quad (7)$$

Finally, Pop-Eleches and Urquiola (2013) identify a specific change in school's inputs (i.e. change in X while Y remains constant)

$$\frac{\partial A_2}{\partial(S_1^x - \bar{S}_1^x)} = \frac{\partial A_2}{\partial S_1^x} = \frac{\partial g_1}{\partial S_1^x} + \frac{\partial g_1}{\partial S_1^y} \frac{\partial S_1^y}{\partial S_1^x} + \frac{\partial g_1}{\partial F_1} \left(\frac{\partial F_1}{\partial(S_1^x - \bar{S}_1^x)} + \frac{\partial F_1}{\partial(S_1^y - \bar{S}_1^y)} \right) \quad (8)$$

Equation 8 indicates the expected program’s impact on student’s achievement relative to non-cognitive skills.

This conceptual model illustrates some parameters that this study aims to identify. In particular, it is expected that the program will generate effects on the outcomes as well as behavioral responses. The educational outcomes are SAT-like test registration and school dropout rates. In addition, life skills, which are proxied using psychometric tests, as well as emotional regulation, which are estimated from neurophysiological recordings, are considered outcomes as well. The student’s behavioral responses are proxied using emotional responsiveness measured through neurophysiological recordings. As mentioned, the interpretation of the behavioral response differs from used by [Todd and Wolpin \(2003\)](#) and [Pop-Eleches and Urquiola \(2013\)](#), because here I only consider the individual behavioral response.

From a policy viewpoint, it is important to understand how life skills can be fostered by young people, especially for those at-risk. In order to fully understand the impact of a public policy it is necessary to study potential behavioral responses.²⁴ In the case of students who participate in a program to promote life skills, it is reasonable to expect behavioral changes in all different life spheres, such as self-perception, emotional regulation, interactions with peers, and behaviors at school and home. Likewise, behavioral responses appear to be context dependent, and thus vary according to each public policy and to changes over time ([Pop-Eleches and Urquiola, 2013](#)).

2.4 Definition and Measurement of Skills

2.4.1 Skill Definitions

[Almlund et al. \(2011\)](#) and [Heckman and Kautz \(2012\)](#) have recently surveyed different approaches and definitions of cognitive and non-cognitive skills. [Gintis \(1971\)](#) defines

²⁴ For example, [Pop-Eleches and Urquiola \(2013\)](#) argue that parents might react to their children going to a better school by lowering their own efforts; or that at-risk students who make it into better schools might feel inferior or stigmatized beside their more privileged peers. Regardless of the relevance of these behavioral responses, developments in the literature have only happened recently (e.g. [Cullen et al., 2006](#); [Pop-Eleches and Urquiola, 2013](#)).

cognitive skills as the process of logically combining, analyzing, interpreting and applying informational symbols. Common examples of these types of skills are the use of language, memory and logical-mathematical knowledge. Non-cognitive skills, also known as socio-emotional skills, life skills or personality traits are harder to define, as they are a combination of behaviors, feelings, values and thoughts.²⁵ Examples of these skills are self-confidence, self-control, perseverance, grit, and locus control. A large body of evidence reviewed in [Almlund et al. \(2011\)](#) suggests that stable character skills exist and are predictive of many behaviors. Moreover, [Borghans et al. \(2008\)](#) highlight the existence of quasi-cognitive skills which are a mix between both cognitive and non-cognitive skills, such as creativity or innovation, emotional intelligence, and practical intelligence, among others. The definition and measurement of cognitive and non-cognitive skills are challenging. Any type of skill is difficult to measure, even when identifiable or separable, but it is even more difficult to define and separate non-cognitive skills from cognitive ones ([Heckman and Kautz, 2012](#)).

Consistent with the labor economics literature, the present study uses self-reported psychometric tests to measure skills. The battery of tests includes psychometric tests that intended to measure non-cognitive, creative and cognitive skills. Regarding socio-emotional skills, the Grit Scale ([Duckworth et al., 2007](#)), Locus Control Scale ([Rotter, 1966](#)) and Big Five Inventory (BFI, [John and Srivastava, 1999](#)) were included in the battery of tests. The Big Five Inventory is a widely accepted taxonomy of personality traits that describes personality as consisting of the following five traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism ([John and Srivastava, 1999](#)). The Grit Scale delves deeper into the measures of the Big Five’s Conscientiousness, defined as “perseverance and passion for long-term goals.” The Locus Control Scale measures the degree to which individuals attribute their fate to external or internal factors. There is an open debate about the malleability of personality in the medium term²⁶, and thus the analysis on the present study will be focused on Grit and Locus Control scales instead of the Big Five Inventory (BFI).

In order to test creativity, the written version of the 1976 Torrance’s Test of Creative Thinking [TTCT] was also used ([Kim et al., 2013](#); [Kaufman and Sternberg, 2010](#)). It is worth mentioning that in the psychology literature on creativity, a distinction is regularly made between convergent thinking, which is referred to here as

²⁵ See for example, [Borghans et al. \(2008\)](#); [Cunha et al. \(2010\)](#); [Almlund et al. \(2011\)](#).

²⁶ See for example [Almlund et al. \(2011\)](#).

cognitive ability, and divergent thinking, which is denoted here as creative thinking.²⁷ The written form of the TTCT, is an assessment tool of creativity that was adapted to Chile by [Egana-delSol \(2016a\)](#). Following [Torrance \(1966\)](#), I shaped an instrument to measure four dimensions of creativity: fluency, flexibility, originality and the “13 creative forces.”²⁸

With respect to cognitive skills, Raven’s Progressive Matrices ([Raven, 1936](#)), which measures fluid intelligence, was considered. The psychology literature usually makes a distinction between crystallized and fluid intelligence ([Almlund et al., 2011](#)). The former accounts for the ability to use past appropriation of the collective intelligence of the culture, giving one a distinct advantage in solving the problems involved, while the latter is “the ability to perceive complex relations, educe complex correlates, form concepts, develop aids, reason, abstract, and maintain span of immediate apprehension in solving novel problems in which advanced elements of the collective intelligence of the culture were not required for solution” ([Almlund et al., 2011](#)). Since the program should not directly affect fluid intelligence, this test is taken as a placebo.²⁹

2.4.2 Measurements of Non-cognitive Skills

There are different approximations to measure non-cognitive skills. A recent survey features the use of behaviors —drugs use, registered behavior at school, etc.— and peer’s and self-reported measures ([Heckman and Kautz, 2012](#)).

For instance, [Heckman et al. \(2006\)](#) use risky behavior —i.e. marijuana consumption during adolescence— to predict later outcomes in the labor market. [Hirschi and Gottfredson \(1993\)](#) argue that objective behavioral measures might be preferred to self-reports, as filling out a survey requires some level of self-control. In addition, answering survey questions is another task that relies on skills beyond those targeted by the survey. Nevertheless, the measurement of risky and irresponsible behaviors in adolescents, such as marijuana consumption or theft, may be related to other context dependent factors rather than socio-emotional skills.

²⁷ For a discussion see [Runco \(2010\)](#).

²⁸ The last dimension comprises of an aggregate and richer assessment of creativity. Unusual visualization, internal visualization, extending or breaking boundaries, humor, richness of imagery, and colorfulness of imagery are some of the dimensions considered. See [Egana-delSol \(2016a\)](#) for details.

²⁹See the Appendix for details.

Using similar measures from administrative data, [Jackson \(2014\)](#) investigates the effect of teachers on students' cognitive and non-cognitive skills. In particular, the author proxies cognitive skills using achievement test scores and non-cognitive skills using absences, suspensions, grades, and grade progression. These measures of non-cognitive skills predict adult outcomes with a strength similar to measures of cognitive ability. His measures of character are commonly available from schools' administrative records. Some scholars criticize this approach, arguing that it is tautological to use measures of behavior to predict other behavior, even though the measures are taken early in life to predict later life behaviors ([Heckman and Kautz, 2012](#)).³⁰ Moreover, it is not always feasible to have access to administrative—or third party—data in a field experiment context, which frequently is nominated and, thus, has restricted access.

Another option is to use others' ratings. However, behaviors reported by peers, teachers or parents could be biased, especially if they know these reported behaviors are used to make decisions that could affect themselves. For instance, a school might like to assess the degree of grit among their students by asking teachers to rate them. Teachers can infer that the results of that assessment could influence their own evaluations, incentivizing teachers to over-rate particular skills in their students.

Moreover, self-reported tests' performance can be influenced by the test setting ([ONeil et al., 2014](#)), reference bias ([West et al., 2014](#); [Kautz et al., 2014](#)), and the examinee's emotional and motivational state ([Plucker and Makel, 2010](#); [Querengsser and Schindler, 2014](#); [Egana-delSol, 2016b](#)). In the context of well designed assessment, it is plausible to argue that the test setting is, on average, similar to everyone. Reference bias is one of the most common caveat for self-reported measures of non-cognitive skills. For example, in the Grit Scale statement "I am a hard worker," the assessment ranges from "Very much like me" (1) to "Not like me at all" (5).³¹ Thus, a subject's reference group will determine their self-perception—a student from a high achieving school might think herself lazy whereas a similar student in a low achieving school may consider herself rigorous.³² Finally, as discussed in [Egana-delSol \(2016b\)](#) emotional/motivational state also affects self-reported measures of non-cognitive skills.

In summary, self-reported tests are problematic instrument to measure non-cognitive

³⁰ For example, a meta-analysis by [Pratt and Cullen \(2000\)](#) finds that behavioral measures are at least as good at predicting crime as are measures based on self-reported taxonomies.

³¹ See [Duckworth et al. \(2007\)](#) for details.

³² For a recent discussion about how reference bias can affect self-reported on non-cognitive skills measures see [West et al. \(2015\)](#).

skills. In response to these caveats, I suggest an alternative method to assess non-cognitive skills, in particular, emotional regulation. In the following section I will explain the relevance of emotional regulation as a non-cognitive skills as well as methods to proxy it eluding self-reporting or past behaviors.³³

3 The Role of Emotions

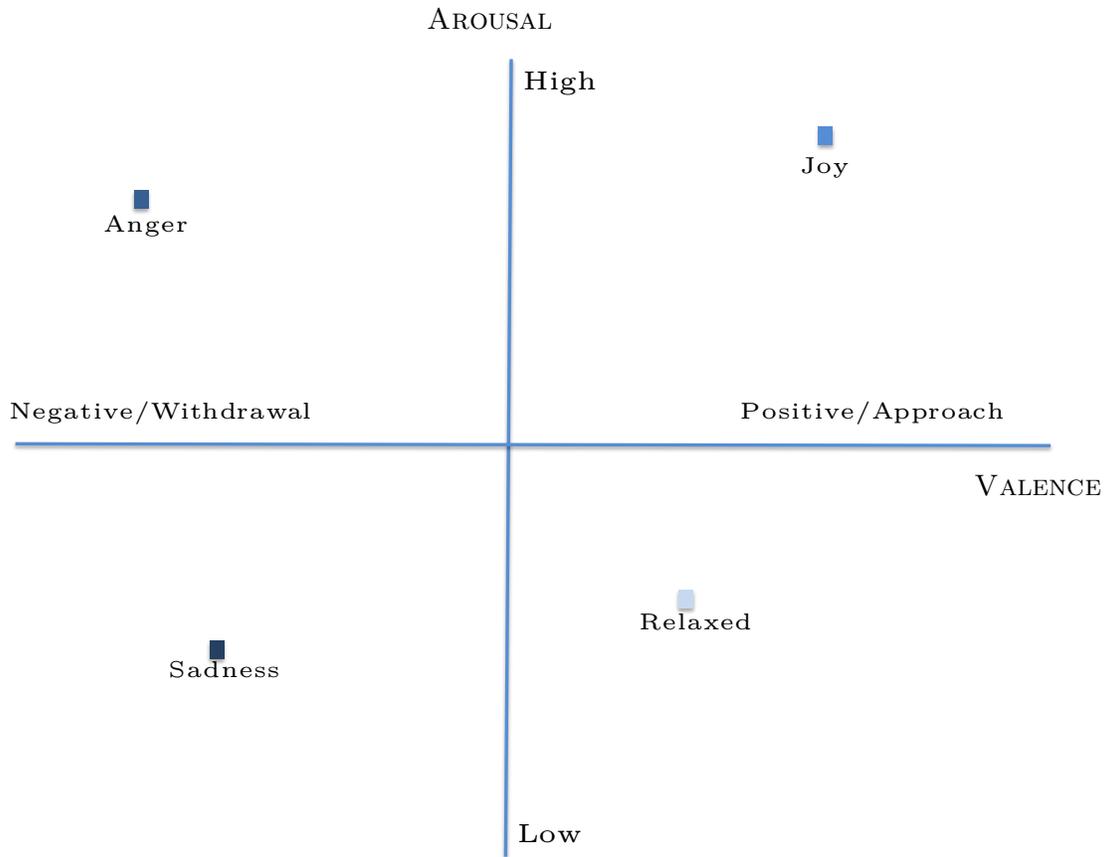
The relation between emotion and cognition has been debated for centuries (Descartes, 1649[1989]). Prominent neuroscientists argue about the critical role that emotions play in cognition, perception, attention and memory (Damasio, 1994; Lakoff, 2008; Salzman and Fusi, 2010; Fuster, 2013). Furthermore, scientists who combine neuroscience, behavioral science, and economics also support the notion that emotions have a strong influence on economic behavior and decision-making, as well as labor market performance (e.g. occupational choice, salaries, entrepreneurship, etc.).³⁴ For example, emotions experienced while making a decision —i.e. choice-option—elicited emotions—immediate emotions are at the base of traditional economic interpretations of utility as emotional carriers of value. Positive emotions increase value and elicit approach, whereas negative emotions decrease value and result in avoidance (Weber and Johnson, 2009). Moreover, emotions unrelated to the judgment or decision at hand, referred as incidental emotions, have also been shown to influence choice (Weber and Johnson, 2009).

Emotions have interesting behavioral implications. For instance, Durlak et al. (2011) conducted a meta-analysis on programs to develop emotional intelligence in educational settings. They found positive effects on targeted socio-emotional skills and attitudes about self, others, and school. They also argue that those types of programs foster students' behavioral adjustment in the form of increased prosocial behaviors, reduced conduct and internalizing problems, and improved academic performance on achievement tests and grades. Similarly, a preschool and early primary school program to foster emotional intelligence improves classroom behavior as well as executive function, defined as higher level cognitive skills including inhibitory control, working

³³ As mentioned, in addition, I explore strategies to correct self-reported psychometric tests from emotional bias using methods from affective neuroscience in Egana-delSol (2016b).

³⁴ See for example Weber and Johnson (2009); Lerner et al. (2012); Lempert and Phelps (2014); Loewenstein (2000).

Figure 1: Arousal-Valence Model of Emotions



memory, and cognitive flexibility (Almlund et al., 2011; Barnett, 2008). In fact, educational or labor market programs that target socio-emotional skills are likely affecting emotional regulation. Complementarily, Haushofer and Fehr (2014) argue that poverty may have particular psychological consequences that can lead to economic behaviors making it difficult to escape poverty. The evidence indicates that poverty causes stress and negative emotional states, which in turn may lead to short-sighted and risk-averse decision-making, possibly by limiting attention and favoring habitual behaviors at the expense of goal-directed ones. Together, these relationships may constitute a feedback loop that contributes to the perpetuation of poverty, which they called a “psychological poverty trap.”

A variety of emotion models are used in affective neuroscience (Petranonakis

and Hadjileontiadis, 2010). The most frequent model used in psychology and affective neuroscience is continuous in nature and thus expresses emotions in a n-dimensional space, usually the two dimension arousal and valence model. In general, the literature uses James-Lange’s arousal-valence —also known as *Circumplex*— model of emotions, which will be followed here. As mentioned, valence accounts for judging whether a situation is positive or negative, while arousal expresses the degree of one’s excitement, spanning from calmness to excitement. In particular, the Ramirez and Vamvakousis (2012)’s neurophysiological version of the model will be considered as the framework for emotions. As an illustrative example, emotions could be classified into four categories in the valence-arousal locus, namely joy, anger, relaxation, and sadness. As Figure 1 reveals, joy is characterized by high arousal and high valence, anger by high arousal and low valence, relaxation by low arousal and high valence, and sadness by low arousal and low valence. Another approach, which is used relatively less frequently by affective scientists but is relevant to neuroeconomics, classifies different emotions, in particular valence, according to motivation (Davidson et al., 1990; Harmon-Jones et al., 2010; Lempert and Phelps, 2014). In this model, different emotional states lead to different goals for action. Positive (valence) emotional states in term of valence, such as happiness, evoke a motive or goal to approach a situation. On the contrary, negative (valence) emotional states such as sadness or disgust, are withdrawal emotions, evoke a motive or goal to withdraw from situations/stimuli linked to these emotions.

Regardless of whether approach/withdrawal motivations correlate with valence, researchers have argued that they represent a distinct dimension (Kassam et al., 2013). For instance, angered individuals are typically motivated to approach the source of their anger —i.e. to fight— despite the stimulus’ negative valence. Nevertheless, both arousal/valence and approach/withdrawal models have been supported by research. They may capture distinct features of human function and may be better thought of as complementary or as simply different (e.g., Coombes et al., 2007; Wacker et al., 2003). The psychological phenomena that these models aim to explain allow for that ambiguity. In particular, the arousal/valence model focuses on emotional stimuli and information processing, rather than behavioral insights, whereas the approach/withdrawal model emphasizes neural activity associated with goal-related emotion, and thus actions, such as approaching or withdrawing from a situation (Spielberg et al., 2008). In summary, I will consider indistinctly positive valence/approach as well as negative valence/withdrawal emotions/motivations, and I will discuss the results considering the

scope of both models.

The complex task of measuring emotions can be done using self-reported measures. These methods generally follow the Positive and Negative Affect Schedule (PANAS, [Watson et al., 1988](#)). Since the PANAS is self-reported, it experiences similar issues as those mentioned above for self-reported psychometric tests. In particular, performance can be influenced by the examinee’s emotional and motivational state ([Plucker and Makel, 2010](#); [Querengsser and Schindler, 2014](#); [Egana-delSol, 2016b](#)), the test setting ([ONeil et al., 2014](#)), and reference bias ([West et al., 2014](#); [Kautz et al., 2014](#)). In fact, emotions may play a role in the measurement error of self-reported psychometric tests ([Querengsser and Schindler, 2014](#); [Egana-delSol, 2016b](#)). Therefore, an alternative can be to use neurophysiological methods from affective neuroscience.³⁵ Emotional regulation —pre-test emotional state and emotional responsiveness to positive and negative stimuli in this study— is the relevant feature respect to the impact evaluation of the mentioned educational program. The neurophysiological methods from affective neuroscience are discussed in next section.

3.1 Emotion-Detection Theory

Over the past decade, emotion-detection research has employed an array of different physiological measurements and methods, including pupil dilation, heart rate, and skin conductance for arousal, and voice and facial manifestations for valence ([Takahashi et al., 2004](#); [Partala et al., 2000](#); [Brown et al., 2011](#); [Verma and Tiwary, 2014](#); [Bos, 2006](#); [Petrantonakis and Hadjileontiadis, 2010](#); [Ramirez and Vamvakousis, 2012](#); [Yoon and Chung, 2013](#)). But many of these measurements can be consciously modified, and thus the signals they produce are not purely objective ([Partala et al., 2000](#); [Ramirez and Vamvakousis, 2012](#)), in particular those related to valence which is the main feature of interest in this study.³⁶ Moreover, arousal effects on behavior, self-reporting and decision-making is ambiguous ([Weber and Johnson, 2009](#); [Egana-delSol, 2016b](#)).

³⁵ As mentioned, I explore strategies to correct self-reported psychometric tests using neurophysiological methods in [Egana-delSol \(2016b\)](#). I argue that valence has a positive correlation: Positive/approach emotions (pleased, happy) imply self-reports are overestimated, while negative/withdrawal emotions (unhappy, sad) would underestimated the self-perception of certain socio-emotional skill. However, for arousal the effect is ambiguous ([Weber and Johnson, 2009](#)).

³⁶ [Ekman et al. \(1987\)](#) suggested the universality of six facial expressions based on Darwinian theory: happiness, surprise, anger, disgust, sadness, and fear ([Darwin, 1872\[2002\]](#)). However, [Partala et al. \(2000\)](#) argue that facials expressions can be easily simulated by subjects.

Electroencephalogram (EEG) recordings allow us to measure brain activity and predict emotional state and physiological responsiveness, improving both accuracy and objectiveness with respect to the aforementioned physiological measurement.³⁷

The firing of neurons in the brain triggers voltage changes. The electrodes in an EEG headset capture the electrical activity corresponding to field potentials resulting from the combined activity of many individual neuronal cells in the brain cortex.³⁸ However, cortical activity measures are distorted by the tissue and skull between the electrodes and the neurons. This introduces noise and reduces the intensity of the recorded signals. Regardless, EEG measurements offer important insights into the electrical activity of the cortex (Verma and Tiwary, 2014; Brown et al., 2011).

A number of authors have considered diverse methods to recognize emotions from EEG recordings, improving both accuracy and objectiveness relative to the aforementioned physiological measurement (Choppin, 2000; Takahashi et al., 2004; Lin et al., 2010; Ramirez and Vamvakousis, 2012; Kim et al., 2013). Accuracy accounts for the true and consistent results, that is to say, the likelihood that the model can correctly predict the elicited emotion, which is known by the characteristic of the stimuli or a self-reported emotional state after a given experiment. Numerous studies have attempted to compare those diverse methods of emotion recognition arguing that the most suitable methodology and strategy —i.e. the one with the highest predictive accuracy— is the use of EEG recordings (Choppin, 2000; Takahashi et al., 2004; Lin et al., 2010; Ramirez and Vamvakousis, 2012; Verma and Tiwary, 2014; Kim et al., 2013).³⁹ For instance, Brown et al. (2011) estimate an 82% accuracy for arousal and valence. Yoon and Chung (2013) found a 70% accuracy for arousal and valence. Verma and Tiwary (2014) found 85% accuracy for arousal, valence and dominance.⁴⁰

Furthermore, the “proof-of-concept” experiment that I conducted on graduate

³⁷ See for example Partala et al. (2000); Brown et al. (2011); Verma and Tiwary (2014); Bos (2006); Petrantonakis and Hadjileontiadis (2010); Ramirez and Vamvakousis (2012); Yoon and Chung (2013).

³⁸ Electroencephalogram devices measure the voltage change that happens when a neuron fires. When a positive change in the voltage crosses a certain threshold, an action potential is triggered. Indeed, the voltage goes from the resting potential of about -60mV to +20mV. This electrical activity from group of neurons is measured in the cortex by the EEG’s electrodes. That is to say, the EEG measures the brain’s activity through voltage change in groups of neurons that had been fired due an action potential.

³⁹ Verma and Tiwary (2014) propose a *fusion* model that encompass other physiological measures in addition to arousal and valence indices. Regardless, the authors estimations indicate that the *Fusion* and *Circumplex* —arousal-valence— models have similar accuracy.

⁴⁰ The dominance scale ranges from submissive (or without control) to dominant (or in control, empowered).

students at Columbia University exhibited 79% accuracy for valence.⁴¹ The recent development of low cost portable EEG devices offers an unprecedented opportunity to incorporate methodologies from affective neuroscience in social programs evaluation and experiments in the field.⁴² In a recent study, [Martinez-Leon et al. \(2016\)](#) compare the quality of data captured by a professional Biosemi Active II⁴³ and a low cost Emotiv EPOC⁴⁴ headset. The latter is identical to the device used in this dissertation. Their results are based on the comparison of the success rate of a Brain-Computer Interface (BCI) system. Higher precision and less variance are found on low cost Emotiv EPOC headset datasets. Moreover, the authors conclude that the Emotiv EPOC low-cost headset can be used on motor imagery BCI systems.

In fact, a suitable strategy to proxy emotional regulation —i.e. emotional state and responsiveness— for dimensions of arousal and positive (approach) and negative (withdrawal) valence (motivation) in the field can be the use of low cost EEG recordings ([Friedman et al., 2015](#)). It is important to note that these indices are not intended to describe and/or predict personality traits or character themselves.⁴⁵

Matlab 2014b software (Waltham, MA, USA), and EEGLAB open source toolbox ([Delorme and Makeig, 2004](#)), were used to do the neurophysiological offline data analysis. The EEG signal was passed through a low-pass filter with a 40-Hz cutoff frequency in order to remove noise coming from the power line and artifacts. The frequency of EEG measurements ranges from 1 to 80Hz, with amplitudes of 10 to 100 microvolts ([Ramirez and Vamvakousis, 2012](#)). Multi-taper Fourier transform for continuous data sets is used. By using a set of tapers, rather than a unique data taper or spectral window, the algorithm reduces the variance of spectral estimate. Moreover, it is particularly effective for short data segments ([Castellanos and Makaro, 2006](#)).⁴⁶ Many authors suggest that a maximum of approximately a 10 second signal would be required if the dependent measure was an EEG, since a longer period can include factors that

⁴¹See [Egana-delSol \(2016b\)](#) for details.

⁴² In a lab setting, [Ramirez and Vamvakousis \(2012\)](#) use the same low cost portable EEG device used in the present study.

⁴³ Biosemi. For details visit <http://www.biosemi.com/>.

⁴⁴ Emotiv EPOC EEG. For details visit <http://www.emotiv.com/>.

⁴⁵ In a recent study, [Korjus et al. \(2015\)](#) show that there is no correlation between resting state EEG waves and any of the five personality dimensions of the self-reported Big Five Inventory. They conclude that their results indicate that the extraction of personality traits from the power spectra of resting state EEG is extremely noisy, if it is even possible. That study does not tackle the mentioned issues around self-reported psychometric tests.

⁴⁶ A MatLab toolbox developed by [Mitra and Bokil \(2008\)](#) to implement those methods is freely available at <http://chronux.org/chronux/>.

are distinct to the elicited emotion (Davidson et al., 1990). Finally, in order to smooth the exponential nature of EEG signals we apply a logarithmic power transformation to the data, which is the standard in this literature (Verma and Tiwary, 2014; Ramirez and Vamvakousis, 2012).⁴⁷

It was decided that the most suitable method to measure emotions from EEG signals is the one implemented by Ramirez and Vamvakousis (2012). Evidence has shown that the prefrontal cortex—in addition to the amygdala and the insula—coordinates consciousness and regulates emotions.⁴⁸ Measuring emotions therefore requires consideration of the EEG signals using electrodes AF3, AF4, F3, and F4 in the 10-20 standard classification, which are located on the prefrontal lobe.⁴⁹ The frequencies of interest in measuring emotions are both alpha (8-12Hz) and beta (12-30Hz) waves.⁵⁰ Alpha waves are predominant in relaxed states and brain inactivation, while beta frequencies are associated with alertness and an excited state of mind.⁵¹ Consequently, the beta-alpha ratio is an effective indicator of the level of arousal. Therefore, arousal index can be estimated by the following equation:

$$arousal_i = \frac{\beta_{F,i}}{\alpha_{F,i}} \quad (9)$$

where F indicates the simple average of electrodes located on the frontal brain areas AF3, F3, AF4 and F4 in the 10-20 universal system,⁵² while i index an individual in the sample. On the other hand, previous neurophysiological evidence had shown that emotional valence has different representation in the right and left brain’s cortical hemisphere.⁵³ Activity decrease over the frontal right region correlates with positive

⁴⁷ The exact transformation follows Aspiras and Asari (2011). In particular, the EEG signals were transformed as follow: $LogSignal_j = 10 * \log_{10}(Signal_j)$; where $j = positive, negative, restingstate$ conditions.

⁴⁸ Bear and Connors (2007). For a recent review see Salzman and Fusi (2010)

⁴⁹ For the exact location of each electrode in the brain, see Figure 10 20 SYSTEM in the Appendix.

⁵⁰ There is ample data linking changes in these frequency bands to various mental processes such as changes in mental state (Moretti et al., 2004), changes in attention allocated to a task (Klimesch, 1999), memory processes (Klimesch, 1996), motivation and emotional processes (Knyazev, 2007), and different sleep stages (Keenan, 1999), among others. Here I focus on arousal and positive valence/approach behavior and negative valence/withdrawal behavior (Davidson et al., 1990; Harmon-Jones et al., 2010; Verma and Tiwary, 2014).

⁵¹ See for example, Bos (2006); Ramirez and Vamvakousis (2012); Verma and Tiwary (2014); Yoon and Chung (2013); Brown et al. (2011); Friedman et al. (2015); Choppin (2000).

⁵² See the Appendix for a diagram with all 10-20 system locations.

⁵³ For a discussion on the validity of estimate valence by comparing hemispherical activation, see, for

emotion, while activity decrease over the left frontal region is related to negative emotion. Since activity decrease —i.e. inactivation— could be measured as the inverse of arousal level, it is possible to estimate valence level by the following relation:

$$valence_i = \frac{\alpha_{F4,i}}{\beta_{F4,i}} - \frac{\alpha_{F3,i}}{\beta_{F3,i}} \quad (10)$$

where $F3$ and $F4$ indicate electrodes located on the left and right frontal brain areas, respectively. Individuals are indexed by i .

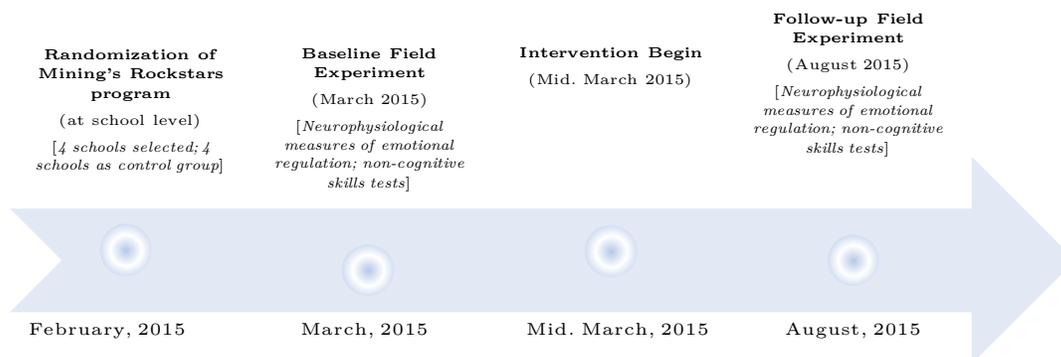
As mentioned before, the literature interprets valence positive and negative indices as approach and withdrawal motivation to stimuli, respectively (Davidson et al., 1990; Harmon-Jones et al., 2010). Evidence from psychology and neurophysiology literature points out that frontal EEG asymmetry is associated with different emotional and psychological states rather than valence. In a seminal work, Davidson et al. (1983) suggested a model called approach/withdrawal theory to investigate frontal EEG asymmetry during emotional states. He claimed that the left pre frontal cortex (PFC) activity is involved in a system facilitating approach behavior to appetitive stimuli, while the right PFC activity participates in a system facilitating withdrawal behavior from aversive stimuli. This model claims that processing related to emotional valence itself is not lateralized in PFC. Rather, emotion-related lateralization is observed because emotions contain approach and/or withdrawal components.

That is to say, frontal EEG asymmetry is perceived because emotions contain approach and/or withdrawal components. Thus, emotions will coexist with either right or left asymmetry, whether they are accompanied by approach or withdrawal behavior (Davidson et al., 1983, 1990; Harmon-Jones et al., 2010; Harmon-Jones and Gable, 2008). Approach/withdrawal motivational states have frequently been linked to asymmetries in left/right frontal cortical activation, especially using EEG, though meta-analyses of fMRI data have failed to find consistent localizations (Kassam et al., 2013).

Finally, there are a number of papers that consider the left, relative to right, frontal cortical activity (LFA) to build behavioral indices of approach or motivation. For instance, Hughes et al. (2014) examines the relation between LFA and effort expenditure

example, Ramirez and Vamvakousis (2012); Kim et al. (2013); Verma and Tiwary (2014); Bos (2006); Harmon-Jones et al. (2010); Yoon and Chung (2013); Friedman et al. (2015); Davidson et al. (1990)

Figure 2: Field Experiment Timeline



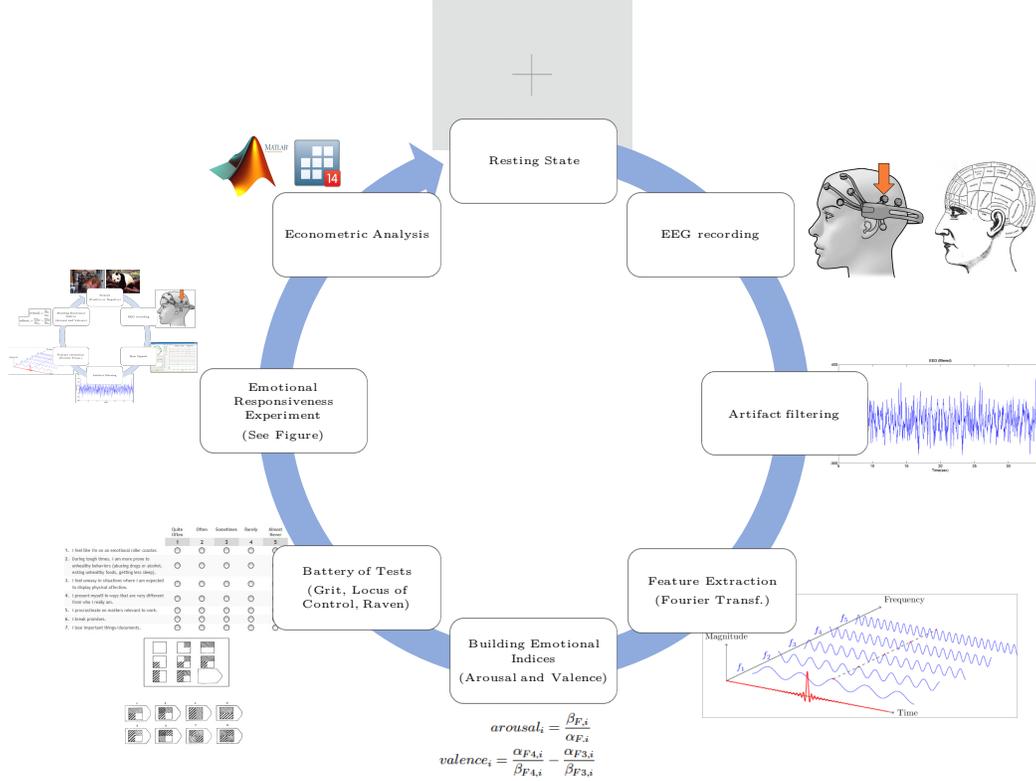
for reward, a behavioral index of approach motivation. They found that subjects with greater resting LFA were more willing to expend greater effort in the pursuit of larger rewards, particularly when reward delivery was less likely.

4 Experimental Paradigm

As mentioned, there exists a common puzzle in the labor economics literature around social programs —i.e. educational or labor market programs— that exhibit positive impacts on labor market outcomes, yet, contrary to expectations, do not seem to affect measures of non-cognitive skills (Card et al., 2010; Calero et al., 2014).

This paper studies the production function of non-cognitive skills —i.e. socio-emotional skills, creativity, etc.— in formal schools. In particular, I will explore how participants in the “Mining’s Rockstars” program change in terms of creativity, grit, locus control, and emotion regulation. I argue that the program affects participants’ emotional regulation —i.e. emotional state and responsiveness. In Egana-delSol (2016b), I

Figure 3: Field Experiment Paradigm



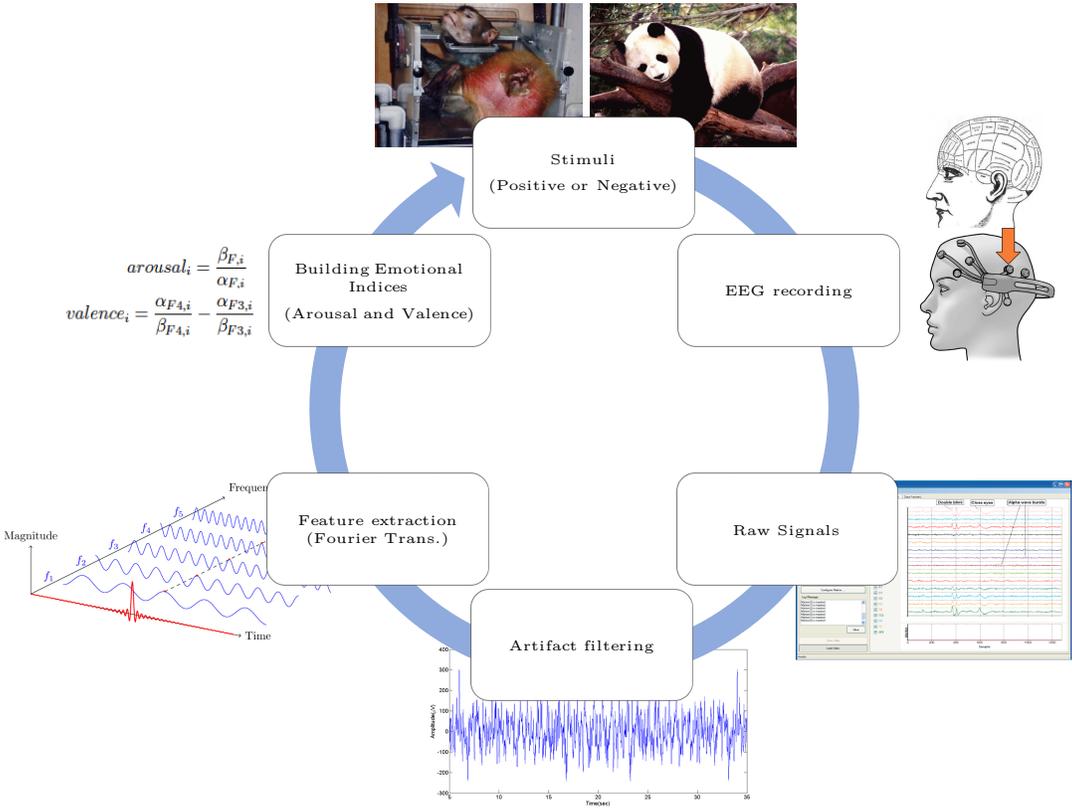
claim that the impact on emotional regulation modulates their self-reporting on tests about their own non-cognitive skills, and thus might be an explanation of the aforementioned puzzle.

As Figure 2 shows, the experiment was done in the context of the impact evaluation of the “Mining’s Rockstars” program. It targeted 1,380 students aged 17-18 years old at baseline in 8 public technical high schools in semi-urban towns in the north of Chile. Four schools were randomly selected to be in the treatment group, while the remaining four schools were assigned to the control group. In the treatment schools, all students in their 4th grade —i.e last year— of high-school were treated.

As mentioned, the typical self-reported method to evaluate emotions is the Positive and Negative Affect Schedule (PANAS). Besides its discussed caveats because of self-reporting,⁵⁴ using PANAS makes salient the relevance of emotions in the study,

⁵⁴See section 1.2 for details.

Figure 4: Experimental Paradigm of Emotional Responsiveness



potentially affecting both reaction to stimuli and self-reports on emotional state and other dimensions. Therefore, I do not use PANAS in this study.

In particular, I collect three streams of data, namely: pre-test resting emotional state from EEG recordings, a battery of psychometric tests (see previous section for details), and emotional responsiveness to both positive and negative stimuli. The pre-test emotional state was constructed using EEG scalp recordings while students watched a black cross in the center of a gray screen for a period of 30 seconds before taking the battery of psychometric tests (see Figure 3). Emotional arousal and valence indices at a resting state are estimated using those recordings. Hereafter, these measures will be called “baseline.” Then, I applied a battery of psychometric tests that includes grit scale, locus control scale, raven-like progressive matrixes, and torrance’s test of creative thinking. The emotional responsiveness measures were obtained right after the students finished the battery of tests. As Figure 4 exposes, the experiment consisted

of showing an alternating series of positive and negative images in order to elicit emotional responses. Arousal and valence indices were also obtained from those emotional responses.⁵⁵ Consequently, there are three estimates of both arousal and valence according to the nature —or absence— of stimuli, including: pre-test resting state, positive and negative.

The final sample —after filtering the EEG data and accounting for attrition— is an unbalanced panel with a baseline in March and a follow up in August 2015, with a total of around 300 valid EEG recordings of students. This is an extraordinary number of measures, considering the time and complexity of collecting this data, especially in an out-of-the-lab setting.⁵⁶

Emotional regulation —both emotional state and responsiveness— are innovative outcomes that try to capture the substantial qualitative evidence that many educational programs or ALMP designed to improve education and labor market outcomes through fostering socio-emotional skills do in fact affect them.

5 Results

This section has four subsections. First, descriptive statistics are shown. Second, the degree of attrition is presented. Third, the program’s impact on educational outcomes —i.e. SAT-like test registration and school dropout rates— as well as non-cognitive skills is estimated. The latter considers creativity and life skills measures, which are the main expected mechanism accounting for program’s impact. Finally, I estimate the program’s effects on emotional regulation, in particular, on pre-test resting state emotional state and emotional responsiveness to both positive and negative stimuli.

5.1 Descriptive Statistics

This subsection presents the descriptive statistics and attrition in the field experiment. Table 1 shows that the sample is balanced at baseline —i.e. March 2014— in all

⁵⁵ Technical details are explained in [Egana-delSol \(2016b\)](#).

⁵⁶ Indeed, I collected almost 500 EEG recordings in students. Filtering for artifacts and data quality gives us the final sample. See Table 2 for details.

Table 1: Experimental Balance at Baseline

	Control	Treated	(1) vs. (2)	p-value dif.
Locus Control Scale	-0.058 (0.070)	0.172 (0.165)	-0.230 (0.166)	0.207
Grit Scale	-0.018 (0.151)	0.294 (0.129)	-0.312 (0.200)	0.162
Creativity Index	0.006 (0.135)	-0.032 (0.155)	0.039 (0.212)	0.861
Creativity Index	-0.083 (0.131)	-0.137 (0.130)	0.054 (0.197)	0.790
ValenceNegative/Withdrawal	0.051 (0.093)	0.107 (0.090)	-0.056 (0.139)	0.699
ValencePositive/Approach	-0.041 (0.072)	-0.061 (0.029)	0.020 (0.071)	0.783
Valence Resting State	0.049 (0.037)	0.080 (0.069)	-0.031 (0.055)	0.592
Arousal Resting State	-0.033 (0.129)	0.277 (0.067)	-0.310 (0.157)	0.089
N	104	76	180	

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard Errors are clustered at school level.

variables but arousal index in resting state.⁵⁷ As Deaton (2010) noted, randomized trials are frequently unbalanced in randomized field experiments, even in large experiments such as the RAND Health Insurance Experiment.

That difference in the arousal index at resting state becomes insignificant, driven by a change in that index among participants. I will discuss these differences later in the difference-in-difference estimation results. These simple results indicate that the puzzle that typically arises in impact evaluations related to educational or labor market programs designed to foster non-cognitive skills is likely to happen in the present study as well. Moreover, as I will show later in this section, the program has significant impacts on at least two relevant educational outcomes —i.e SAT-like test registration and school dropout rates.

⁵⁷ As mentioned, the data was collected in August instead of June because of national teacher strike at the end of the first semester. Then, many schools prefer to begin the follow up collection of data in August instead of the very beginning of the second semester —i.e. middle of July— because they already had many activities scheduled.

5.2 Attrition

It is important to analyze data attrition —i.e. missing data— in experimental studies. The main goal is to understand its sources and test whether the attrition was uneven across experimental groups. There are different sources of data attrition in this study.

First, experimental attrition generated in the baseline due the lack of precision in the collected metadata that was later used to merge different sources of information. Field experiments were implemented using several working stations —i.e. a laptop with a Emotiv EPOC headset. There were six and eight stations in the baseline and follow up, respectively. When conducting the experiment, we made several mistakes linking metadata on the computers, written records to match individuals with their data files, specially when an error happened and it was necessary to restart the whole experiment. In particular, it was necessary to record the initial and final time of each individual. With that information it was possible to match the EEG-data files with test scores.

Second, there was attrition as a result of the quality of the recording. Due to different issues, EEGLAB —the Matlab toolbox used to analyze EEG data— sometimes failed to read the EEG recordings properly. Highly dense amounts of hair, long hair, or both, and/or computers freezing during experiments were the usual suspects.⁵⁸ Moreover, more subjects were rejected after signal processing —high and low band-pass filtering, Fourier transformation into frequencies, etc. In short, many subjects were lost.

Table 2: Attrition

	Baseline			Follow up		
	Total	Treated	Control	Total	Treated	Control
Participants	287	123	164	205	113	92
Valid EEG Records	195	85	110	136	68	68
Valid EEG w/relevant info.	180	76	104	116	64	52

Notes: “Valid EEG records” indicates subjects who have EEG data that is readable. “Valid EEG w/relevant info.” accounts for subjects with readable EEG data and that was also possible to match it with test scores.

⁵⁸ For example, Laptop number 5 recorded with a high proportion of error for two consecutive days, which generate 13 missing subjects.

Third, individual missed between experiments rounds —i.e. between baseline and follow up field experiments— can be an important source of selection bias due attrition. There was a decrease in the total sample size of around 36%. Moreover, it is important to highlight that the attrition in the control group, which rose to 43%, is considerably larger than that in the treatment group. Nevertheless, the attrition in the control group was balanced with respect to the measurements considered in this study. In fact, there is no statistically significant difference at baseline between treatment groups that participated in the experiment only at the baseline and those whom participated at both baseline and follow up.⁵⁹

Fourth, It is also possible to rule out some potential unobservables characteristics that could be present in a major proportion of the group of students who participated in the follow up. For example, the degree of intrinsic motivation could be different, which is allegedly a proxy in the locus control test, and did not show differences in the baseline (see Table 1). Anecdotically, references from teachers at the control schools explained some of the reasons behind the lack of participation in the follow up. For one thing, many students have a lack of motivation in general, especially to attend school. In fact, many of the control schools' students who participated in the baseline experiment were absent from school that day. There were also a couple of cases of students who dropped out of school altogether.

Fifth, it is plausible that boredom contributed to the lack of incentive to participate in the study, especially in the control group which has no relation to the program. Students in the control group were completely uninformed about what and why we were doing this study on them. Participating in the baseline experiment could be boring. The experiment takes 30-40 minutes on average, requiring students to be seated in the classroom conditioned as a laboratory to set up the EEG headset, check and habituate the EEG headset to have a correct reading of all 14 electrodes, and then take the actual test, which is 20 minutes on average. Indeed, the relative large number of takers —not considering whether they have a valid EEG datafile, see Table 2— in the control group in baseline (164 subjects) was unexpectedly high compared with those in the treatment group (123).

Potential selection bias due to attrition is unlikely because it happened apparently randomly in both treatment and control groups. See the the Table 7 in the appendixes for the mean test of the difference of observable characteristics at baseline between

⁵⁹See Table 7 in the appendixes.

participant and control group’s students missed between the baseline and the follow up, and those who had the change to do follow up as well.

5.3 Impact on Educational Outcomes

The Mining’s Rockstar program is expected to affect some educational outcomes. Here I consider two outcomes —i.e. SAT-like⁶⁰ registration and dropouts rates— in order to argue that the program has observable impacts on the treatment group, aside from its impact on the emotional regulation indicators.⁶¹ This is particularly important since it shows that this program is experiencing a similar puzzle to the one found on the labor economics literature. That is to say, on the one hand, the program exhibits a positive impact on educational outcomes —i.e. SAT-like registration and dropouts rates. On the other hand, as shown in Table 6, the expected factors behind the impact —i.e. non-cognitive skills— are not affected.

The administrative data on educational outcomes is only accessible at the end of the academic year. Thus, I estimate a simple difference model. In addition, since the data comes from administrative records, it was possible to consider additional control groups as a robustness check of the results. The model is the following:

$$Y_{j,i,t} = \alpha + \beta * treat_i + \theta * X_{i,t} + \varepsilon_{i,t}, \quad (11)$$

where $Y_{j,i,t}$ indicates the j educational outcome of individual i in time t , with $j = SATregistration, dropoutrates$, $i = students$, and $t = follow - up$. Furthermore, $treat_{j,i}$ indicate who had been treated by the program. Finally, $X_{i,t}$ indicates school level dummy variables.

The program had a significant impact on educational outcomes —i.e. dropout rates and SAT-like registration; see Table 3 and 4, respectively.— yet had no impact

⁶⁰The University Admission Test (Spanish acronym PSU) aims is to measure the mastery of secondary education topics. It is similar to the SAT in the US. This test is voluntary for students who have finished the secondary education, but required for applying to university. The PSU is the principal selection mechanism along with average secondary education grades, which are combined. The PSU, has a relative weight in the final university application score of around 30%. The PSU score is normalized to an average 500 points and a standard deviation of 100 points, with a fixed status ranging from 150 to 850 points. Thus the PSU and secondary school grades are the principal instruments to access the most prestigious careers and universities in the country.

⁶¹See next subsection for details about the program’s impact on emotional regulation.

on socio-emotional skills or creativity measures, which is consistent with the findings in the labor economics literature referenced above.

Table 3: Registration to take PSU (SAT-like) test

VARIABLES	(1) Registered	(2) Registered	(3) Registered	(4) Registered
Treatment (RCT’s Schools)	0.130*** (0.0276)	0.221*** (0.0297)		
Treatment (RCT’s EEG Subsample)			0.309*** (0.0556)	0.375*** (0.0585)
Non-Municipal School		0.295*** (0.0353)		0.272*** (0.0865)
Observations	1,296	1,296	294	294
School FE	No	No	No	No

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Notes: School dummies and other controls are only available for registered students. Due to privacy restrictions of the PSU dataset, it was not possible to identify the schools in the sample. However, it was only possible to control for one school that is run by an external non-for-profit institution —instead of the DAEM— where students tend to enroll in the tests at a larger proportion.

As Table 3 shows, there is a significant difference in registration for all models of SAT-like (PSU in Chile) registration. Registration for the SAT is voluntary, but is a requirement to enroll in any university. In order to attend a technical college —i.e. something similar to community college in the US— it is not necessary to take the SAT-like test. In fact, the positive impact found on the treated compared to both control groups in the SAT-like test registration rate can reflect an increase in their expectations about their futures.

Table 4 shows a statistically significant decrease in dropouts rates of about 6 percentage points among participants. This decrease in dropout is a relevant outcome. Dropout rates are high in Chile, especially among students who come from vulnerable socio-economic backgrounds, and who are over-represented in technical schools (Valenzuela et al., 2013). Moreover, descriptive analyses of similar programs implemented previously by the same NGO that runs “Mining’s Rockstars” program, *Emprende Joven*, had shown a positive impact of the program decreasing dropout rates among participants with respect to historical school dropout rates (Albarran and Gonzalez, 2015). In addition to the social desirability of school retention for many reasons —e.g. in

Table 4: Students Dropout Rates

VARIABLES	(1) Dropout	(2) Dropout
Treatment (RCT's Schools)	-0.0683* (0.0366)	
Treatment (Vocational Sch. Same Region)		-0.0428*** (0.00678)
Sex	-0.00228 (0.00968)	0.000849 (0.00417)
Observations	1,888	8,017
School FE	Yes	Yes

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: “Dropout” is a dichotomous variable that equals to 1 if the student dropped out, 0 otherwise. Without loss of generality, only students in their last year of secondary school were considered controls in the analysis. Marginal effects of Probit model reported.

Chile it is mandatory to finish high school— the negative impact found is consistent with the positive impact on expectations. In summary, it is important to support the effectiveness of the program to affect observable educational outcomes. However, it is beyond the scope of this study to discuss further the impacts on educational outcomes because the focus is on the mechanisms, such as the program’s effects on emotional regulation or non-cognitive skills.

5.4 Impact on Self-reported measures of Non-cognitive Skills

It is expected that the Mining’s Rockstar program affects non-cognitive skills. In particular, using conventional self-reported tests, grit—perseverance— and locus control were measured.⁶² The program is also planned to impact on creative and innovative skills, which is a mixture between cognitive and non-cognitive skills.⁶³ In addition, measuring creative skills is scarce in the literature, in particular the context of an educational program evaluation.⁶⁴ The model to be estimated will be the following:

⁶²See previous section for details on these tests.

⁶³See [Egana-delSol \(2016b\)](#) and [Egana-delSol \(2016a\)](#) for details on creativity definitions and measures.

⁶⁴See [Egana-delSol \(2016a\)](#) for a review of the literature on creativity and educational programs.

Table 5: Difference Model on Non-cognitive Skills: Creativity, Locus Control and Grit

VARIABLES	(1) Creativity Index	(2) 13 Forces Creativity	(3) Locus Control Test	(4) Grit Test
Treatment	0.170 (0.260)	0.205 (0.208)	-0.113 (0.300)	0.739** (0.270)
Constant	-0.250 (0.164)	-0.144 (0.159)	0.191 (0.241)	-0.150 (0.271)
Observations	116	116	116	116
R-squared	0.181	0.181	0.022	0.147
School Dummies	Yes	Yes	Yes	Yes
Schools's Class Cluster	Yes	Yes	Yes	Yes
4th Grade	Yes	Yes	Yes	Yes

Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard Errors are clustered at individual level. “4th Grade” is dichotomic variable that index if student was in 4th grade or not. This variable was statistically non-significant in all specifications.

$$Y_{j,i,t} = \alpha + \beta * treat_i + \gamma * Post_t + \delta * treat_i * Post_t + \theta * X_{i,t} + \varepsilon_{i,t}, \quad (12)$$

where $Y_{j,i,t}$ indicates the j test’s score of individual i in time t , with $j = creativity, grit, locuscontrol$, $i = students \in FieldExperiment$, and $t = baseline, follow - up$. Furthermore, $treat_{j,i}$ and $Post_{j,t}$ indicate who had been treated by the program and survey time —i.e. baseline or follow up— respectively. Finally, $X_{i,t}$ indicates school level dummy variables and 4th grade dummy variable. In order to have the highest possible statistical power in the estimations, it was considered both non-treated 4th grade (N=61) and 3rd grade (N=95) students as control in all model specifications, controlling by a dichotomic variable to differentiate them.

As Table 1 shows, there is a significant difference in pre-test resting state arousal in baseline measures in favor of the program’s participants.⁶⁵

Furthermore Table 5 estimates the program’s impact on the treatment group, considering only the follow up measures. This implies $t = august$, and thus $\gamma = \delta = 0$. Since the experiment was randomized, the results contained in Table 5 are consistent

⁶⁵The estimation of balance on Table 1 considered standard errors that are clustered at school level. This allows to control for difference between schools at baseline, which is particularly important here since the randomization was done only among eight schools.

and unbiased estimations of the average treatment effect on the treated. In particular, there is a positive and significant impact on Grit Scale of 0.74 standard deviations(σ), which is near the upper bound on similar interventions.⁶⁶

However, in order to rule out differences that could arise in the baseline at individual level, a differences-in-differences model (DiD) is estimated. This kind of model incorporates the differences between the treated and control groups at baseline, before the program started. Moreover, it also contemplates a common time trend for both groups in the outcomes variables of interest.

Table 6 exhibits the estimates from equation 12 of the DiD model. The main parameter of interest is δ that isolates the average treatment effect taking into account the natural evolution of the outcome across time as well as the selection into the treatment group. In fact, Table 6 shows that program participation apparently does not positively impact grit or locus control scales, which is contrary to the expected impact by program design. Similarly, there is no impact on either the creativity or the “13 creative forces” indices. In addition, if we estimate a model for each of the three dimensions of creativity, there is only significant impact on the *originality* dimension.⁶⁷

Consistent with the results on Table 6, the impact on Grit Scale is likely driven by a difference in selection to treatment rather than the program itself. The Locus Control Scale exhibits no significant change, however the coefficient is negative, which indicates a trend towards internal locus control that is positively related to educational and labor outcomes (Puentes and Urzua, 2010). Furthermore, the Torrance Test of Creative Thinking comprises of three dimensions of creativity, plus an additional indicator of “overall” creative skill, namely: fluidity, flexibility, originality and “13 creative forces.” In fact, the estimated program’s impact on the creativity index as well as the “13 creative forces” index are not significant.⁶⁸

⁶⁶For recent reviews see, for example, Murnane and Ganimian (2014); Duflo et al. (2012).

⁶⁷Estimations using either a simple average of the three dimensions of creativity —flexibility, fluidity and originality— or a principal component analysis index (pca) of creativity exhibited similar results in terms of magnitude, sign and significance. The only dimension on which the program had a positive and significant impact was *originality*. These results are available from the author upon request.

⁶⁸Analogous to the DiD model, estimations using either a simple average of the three dimensions of creativity —flexibility, fluidity and originality— or a principal component analysis index (pca) of creativity exhibited similar results in terms of magnitude, sign and significance. The only dimension on which the program had a positive and significant impact was originality, which is also similar to the difference-in-difference model.

Table 6: Difference in Difference Model on Non-Cognitive Skills: Creativity, Locus Control and Grit

VARIABLES	(1) Creativity Index	(2) 13 Forces Creativity	(3) Locus Control Test	(4) Grit Test
Treatment & Post.	0.243 (0.255)	0.172 (0.236)	-0.105 (0.245)	0.0994 (0.203)
Treatment	0.0300 (0.199)	0.0424 (0.201)	0.327 (0.262)	0.706*** (0.220)
Post.	0.0545 (0.167)	0.318* (0.171)	0.166 (0.163)	-0.0179 (0.147)
Constant	-0.0811 (0.222)	-0.179 (0.224)	0.0129 (0.193)	-0.183 (0.166)
Observations	296	296	296	296
R-squared	0.074	0.086	0.029	0.110
School Dummies	Yes	Yes	Yes	Yes
Student Level Cluster	Yes	Yes	Yes	Yes
4th Grade	Yes	Yes	Yes	Yes

Robust Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: Standard Errors are clustered at individual level. “4th Grade” is dichotomic variable that index if student was in 4th grade or not. This variable was statistically non-significant in all specifications.

5.5 Impact on Emotional Regulation

Table 7 describes the program’s impact following a DiD model that considers two novel outcomes attempting to capture the unobservable impact on emotional regulation. First, I use the indices of pre-test resting state valence and arousal following the methodology explained above. Second, I proxy emotional responsiveness to both positive and negative stimuli also in the arousal and valence dimensions. In simple words, emotional responsiveness captures the neurophysiological reactions of students when they view an emotionally laden set of images.⁶⁹

In particular, the following DiD model is estimated:

$$E_{j,i,t} = \alpha + \beta * treat_i + \gamma * Post_t + \delta * treat_i * Post_t + \theta * X_{i,t} + \varepsilon_{i,t} \quad (13)$$

where $E_{j,i,t}$ indicates the j emotional state or emotional responsiveness associ-

⁶⁹ See [Egana-delSol \(2016b\)](#) for examples images from the Geneva Affective Picture Database (GAPED).

Table 7: Impact on Emotional Regulation: Emotional State and Responsiveness

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	ArouRest	ValRest	ValNeg	ValPos	ValNegDif	ValPosDif
Treat & Post.	-0.591** (0.233)	-1.350* (0.802)	0.649 (0.636)	-0.0524 (0.802)	-1.999** (0.954)	-1.298 (1.135)
Treatment	0.0669 (0.196)	0.129 (0.436)	0.0130 (0.592)	0.597 (0.571)	0.116 (0.669)	-0.468 (0.695)
Post.	0.00761 (0.155)	0.255 (0.626)	-1.079** (0.470)	0.468 (0.568)	1.335* (0.736)	-0.213 (0.804)
Constant	-0.0639 (0.190)	0.191 (0.588)	-0.219 (0.774)	-1.319 (0.814)	0.410 (0.857)	1.510 (0.924)
Observations	296	296	296	296	296	296
R-squared	0.059	0.024	0.038	0.016	0.034	0.027
Student FE	Yes	Yes	Yes	Yes	Yes	Yes
Student Clust.	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
4th Grade	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: “4th Grade” is dichotomic variable that index if student was in 4th grade or not. This variable was statistically non-significant in all model specifications.

ated with an emotionally laden stimuli of individual i in time t , with $j = arousal_{base}, valence_{base}, valence_{positive}, valence_{neg}, dif - valence_{pos}, dif - valence_{neg}$, $i = students \in FieldExperiment$, and $t = baseline, follow - up$. Indicators $dif - valence_{pos}$ and $dif - valence_{neg}$ identified the difference for each individual i at time t between her $valence_{base}$ and $valence_{positive}$ and $valence_{neg}$, respectively. That is to say, the indicators capture the emotional reaction considering the resting state level of valence. As before, $treat_{j,i}$ and $Post_{j,t}$ indices who had been treated by the program and survey time—i.e. baseline or follow up— respectively. Finally, $X_{i,t}$ indicates school level dummy variables. It is expected that $\delta < 0$ and $\beta = \gamma = 0$.

As mentioned, the DiD model is a suitable choice considering baseline disparities between treatment and control groups, in particular the difference shown in the pre-test resting state arousal index. Table 7 shows that both participants’ pre-test resting state arousal and valence indices experience a statistically significant decrease compared to their control group. Impacts are 0.14σ and 0.43σ on pre-test resting state arousal and valence, respectively. Moreover, as Table 7 shows, there is a significant impact on responsiveness to negative emotionally laden stimulus— around 0.46σ , which is near the upper bound on similar interventions (Murnane and Ganimian, 2014; Duflo et al.,

2012). In other words, subjects’ neurophysiological responsiveness when faced with negative stimuli —i.e. pictures of mistreated animals or human rights violations— change significantly after treatment compared with their own baseline reaction and the control group emotional reaction. This might be interpreted as an increase resilience trait among participants.⁷⁰

Similarly, I consider the model in differences on the dependent variable following the specification of the equation (14).⁷¹ By doing this, all individual time-invariant unobservable factors that are relevant for the model are cancelled out. It is important to note that here we restrict our analysis to the balanced panel data, which consists of only 68 observations. The significant and negative impact on emotional responsiveness to negative stimulus —Model (8) on Table 10 in the Appendix— is consistent with previous results.

In summary, the impacts found on emotional regulation are statistically relevant compared to those that similar programs have on other outcomes of interest (Murnane and Ganimian, 2014; Duflo et al., 2012). As mentioned, there is extensive literature about decision making and emotions, and also about behavioral change and emotions.⁷² However, the results here are not trivially translated into a behaviorally meaningful interpretation.

For instance, similarly, Ibarraran et al. (2014) argue that impacts from ALMP using self-reported measures of non-cognitive skills, yet statistically significant, have unclear practical and behavioral interpretation respect to outcomes in the labor market.

A discussion of the results and some examples can be found in the next subsection.

5.6 Discussion

Nowadays, evidence supports the argument that cognitive skills explain only a small fraction of the labor market outcome variance (Almlund et al., 2011; OECD, 2015). The literature has been focused on improving the robustness of the estimations of non-cognitive skills’ models using factor models (e.g. Cunha and Heckman, 2008; Attanasio

⁷⁰See for example Troy and Mauss (2011)

⁷¹See details about the specification of this model in the Appendix.

⁷²For recent literature reviews see for example Weber and Johnson (2009); Loewenstein et al. (2001); Lempert and Phelps (2014); Lerner et al. (2012)

et al., 2015), with the idea of recovering a given underlying skill using at least two—or three, depending on the model—different proxy measures of the same skill. This strategy increases the reliability of a skill’s proxies, but does not solve the problem that the program potentially affects other unobservable dimensions such as emotional regulation, and thus the way in which subjects answer self-reported tests. This is particularly important for both surveys and single trial measurements, when it is not possible to have revealed outcomes—i.e. salaries or employment conditions— or when the aim is to understand the channels driving observed outcomes.

I argue that programs meant to enhance life skills likely impact dimensions that are too complex to measure through traditional self-reported psychometric tests. This complexity arises because programs that potentially impact emotional regulation can affect behaviors in many ways, including the actual self-perception of subjects at the moment in which they are answering self-reported tests. Furthermore, I explored the relevance of the development of emotional regulation, which was identified as a pretest arousal and valence indices, as well as emotional responsiveness to both positive and negative stimuli.

In short, I find no significant impact of participation in the program on different measures of non-cognitive skills that the program is supposed to foster, which is consistent with the literature.⁷³ However, I found a significant program impact on pre-test emotional state, measured by a decrease in arousal and valence indices. In fact, the decrease on the indices of emotional state—i.e. arousal and valence—potentially accounts for the aforementioned puzzle in the literature. Moreover, I estimate a decrease in the emotional responsiveness to negative stimuli that is also significant. Considering the statistical significance of these results, and the novelty of the methodology and indices for a social program evaluation, I will offer a discussion with plausible explanations derived from psychology, behavioral economics, and affective neuroscience literature. The discussion will be structured to consider (i) the impact on resting-state emotional state, and (ii) emotional responsiveness to external stimuli.

⁷³See for example Calero et al. (2014); Heckman and Kautz (2012); Kautz et al. (2014).

5.6.1 Resting-state Emotional State

Consistent with previous studies⁷⁴, I found no short-run impact among “Mining’s Rock-stars” participants in the expected outcomes, such as grit or locus control. This apparent lack of impact is consistent with the decrease on both arousal and valence indices generated by program’s participation. Indeed, since there is in general a positive correlation between emotional state and self-reporting on non-cognitive skills tests, non-significant impacts are expected.⁷⁵ At first glance, this negative impact on emotional dispositions seems to have the opposite effect to that of a program aiming to improve socio-emotional skills following a methodology that relies on the principle of learning by failure. The program’s methodology also includes the analysis of failure—which is endogenously generated in every weekly activity— awareness about context, and potential opportunities to improve. The apparent negative impact on emotional state might be theoretically driving the estimated lack of impact on the skills of interest. In the following sections I will review a number of potential interpretations of these counterintuitive results.

Emotions and Awareness

In terms of the empirical findings of the present study, a decrease in pre-test arousal can be completely positive. In particular, the level of arousal was higher at baseline for participants who can be thought to have an altered emotional state due external adverse factors, such as being surrounded by extreme poverty or violence. In fact, a decrease in the pre-test arousal indices can imply a more relaxed state of the individual.

Emotions can be interpreted in many ways. Darwinian emotions, modulated by the reptilian brain, are the most basic. Among these are emotions such as anger, fear of snakes and spiders, etc. The James-Lange theory argues that there is a correlation between physiological experience —i.e. feelings— and emotions. This theory is the basis for the arousal-valence model that I have used here. [Damasio \(1994\)](#) elaborated a similar theory, called the somatic-marker hypothesis. He argues that somatic-sensory feelings are coded —i.e. related— with past experience, which has associated emotions.

⁷⁴See for example [Attanasio et al. 2015](#); [Calero et al. 2014](#); [Card et al. 2011](#)

⁷⁵See [Egana-delSol \(2016b\)](#) for details in the relation between emotions and self-reporting.

In general terms, valence can be understood as attitude and disposition to act.⁷⁶ In other words, it is the departure state of body and mind before taking any action. Indeed, emotions modulate the course of an individual's actions. They can also be seen as the framework for actions. For instance, if an individual is experiencing a low emotional state —i.e. she is low in arousal and valence— it is harder for her to engage in a given task. For example, [Barraza et al. \(2015\)](#) study the relation between emotional state and memory. The authors argue that positive emotions create an orientation towards the outside world. By contrast, a negative emotional state is associated with an orientation towards yourself, to the internal body/mind.

Moreover, low levels of arousal and valence could be thought of as depressive, and thus a more realistic and pessimistic state ([Barraza et al., 2015](#)). This claim is in opposition to those experiencing positive emotions, who are likely optimistic about their own situation and the environment. Moreover, subjects who are emotionally depressed focus their attention relatively more to the exterior, compared to those experiencing positive emotions who tend to focus more on themselves ([Barraza et al., 2015](#)). Therefore, the lower level of arousal and valence can be interpreted as an increase in external awareness. The program likely affects the sense of awareness among participants through its work on failure analysis. For instance, due to the high level of vulnerability of the context of this study, the withdrawal behavior can be thought of as a lower level of frustration because adverse outcomes are expected. Moreover, the results asymmetry respect to emotional responsiveness —i.e. there is significant impact on decreasing reaction only for negative stimulus— are consistent with those found in research about individuals who meditate performing mindfulness meditation. Meditation practitioners are less affected by stimuli with an adverse emotional load, while their processing of positive stimuli remains unaltered ([Sobolewski et al., 2011](#)).

In short, the results obtained in the present study can be interpreted as a change in attitude and disposition to act and the level of awareness among participants. Therefore, we should observe a dynamic complementarity between those changes and other skills acquisition by participants. For instance, changing attitudes towards failure would also change educational and labor market decisions, such as effort given to tasks, occupational choice, entrepreneurship, and the pursuit of creative and original work, among many others. The exploration of those dynamic complementarities is an interesting topic for further research.

⁷⁶See for a recent review [Harmon-Jones et al. \(2010\)](#) or [Salzman and Fusi \(2010\)](#).

Approach/withdrawal Theory of Emotion

As mentioned, [Davidson et al. \(1983\)](#) suggested, in a seminal work, a model called approach/withdrawal theory to investigate frontal EEG asymmetry during emotional states. Since then, psychology and neurophysiology literature points out that frontal EEG asymmetry is associated with different emotional and psychological states in addition to valence ([Davidson et al., 1990](#); [Harmon-Jones and Gable, 2008](#); [Harmon-Jones et al., 2010](#)). In fact, the approach/withdrawal model invites us to think of an additional interpretation of our results. The decrease in the frontal EEG asymmetry —i.e. valence index— are now consistent with a relative increase of withdrawal behavior. Due to the high level of vulnerability in the context of this study, that withdrawal behavior can be thought of as a lower level of frustration of participants relative to the control group because adverse outcomes are expected. Since the Mining’s Rockstars program aims to master that attitude towards failure, this is an appropriate lens through which to interpret our results.

Efficiency in the use of Resources

Finally, an alternative explanation to the puzzle in the literature that motivates this study regards efficiency in the use of resources. I argue that emotional regulation, which modulates the attitude and disposition to act in the external world, is affected by the program. In particular, it is possible to frame the change under the concept of mental state. Mental state is a disposition to action —i.e., every aspect of the individual inner state that can contribute to its behavior or other responses— that is present at a given moment ([Salzman and Fusi, 2010](#)). Since many non-cognitive skills are thought to be non-malleable in the short term, it is possible that the program of study does not directly affect a certain personality trait or skill, but instead affects the degree of efficiency given the subject’s set of cognitive and non-cognitive skills or mental state. That is to say, participants somehow learn to use their personal resources more efficiently, take opportunities, modulate negative shocks, and so forth. These claims fit well into emotion theories that associate positive affects with approach motivation and negative affects with withdrawal motivation ([Harmon-Jones et al., 2010](#)). It is also plausible that the emotional state has a nonlinear relationship resembling an inverted U with respect to performance in cognitive or other behaviors. This would be consistent with the famous Yerkes-Dodson law that states that stress and performance may exhibit a nonlinear relationship resembling an inverted U ([Haushofer and Fehr, 2014](#)). According to the Yerkes-Dodson law, moderate increases in arousal lead to improvements in

performance, whereas extreme levels of arousal lead to performance decrements. Since we observe a decrease in arousal and allegedly positive behavioral impacts, participants should be in the right side of the peak in the inverted U. This can be an interesting topic for further research.

5.6.2 Emotional Responsiveness

The impact on emotional responsiveness is of particular interest. We observe a decrease in the neurophysiological reaction to negative stimuli in participants, while observing no significant impact on the responsiveness to positive stimuli. This result could be phrased as a relevant outcome of the “Mining’s Rockstar” program. This result is consistent with recent evidence of late stage investments showing a positive impact on behaviors instead of on underlying non-cognitive skills (Blattman et al., 2015).

Reappraisal Strategy

A core aspect of the program’s methodology includes analyzing and rethinking failure, which is intentionally generated in each of the weekly activities contained in the program’s guidebook. Indeed, it is possible to frame that part of the methodology as a reappraisal of the emotions involved, especially in the case of failure. Lerner et al. (2015) review the recent findings for “Solutions that Seek to Minimize the Emotional Response.” They point out four dimensions: Time delay, Suppression, Reappraisal, and The “dual-emotion solution” (inducing a counteracting emotional state). In particular, reappraisal consists of reframing the meaning of stimuli that led to an emotional response. Reappraisal has consistently emerged as a superior strategy for dissipating the emotional response (Lerner et al., 2015; Gross et al., 2003). Specifically, reappraisal includes such behaviors as reminding oneself “it’s just a test” after receiving a poor exam grade, adopting the mind-set of a nurse or medical professional to minimize the emotional impact of viewing someone’s injury, or viewing a job layoff as an opportunity to pursue long-forgotten dreams (Gross et al., 2003). In contrast to suppression, reappraisal not only reduces self-reported negative feelings in response to negative events but also mitigates physiological and neural responses to those events (Lerner et al., 2012; Ochsner et al., 2002; Jamieson et al., 2012). Additionally, regulating emotion with reappraisal-focused strategies that encourage taking a different perspective has been shown to reduce loss aversion in decision making (Sokol-Hessner et al., 2012).

That result has been observed both in choices and in the relative arousal responses to actual loss and gain outcomes.

Finally, there is evidence that reappraisal-focused strategies increase resilience trait among participants, which is consistent with the hypothesis that the program is affecting that skill.⁷⁷

Behavioral Corollaries

In behavioral economics literature, even minor mood manipulations have a substantial impact on emotions and behavior (DellaVigna, 2009). For instance, individuals have been found to tip more at restaurants on sunnier days (DellaVigna, 2009). International soccer matches impact the daily stock returns for the losing country by 0.21 percent (Edmans et al., 2007), however there is no significant impact when a country wins. World Cup elimination games have larger effects. Whether or not the loss was expected does not affect the results. On a different topic, Dahl and DellaVigna (2009) estimate the short-run impact of exposure to media violence on violent crime, that can be phrased as a negative emotional stimulus. They find that violent crime is lower on days in which exposure to media violence is higher, which seems counterintuitive with respect to the laboratory evidence. Exposure to movie violence can lower violent behavior relative to the foregone alternative activity —i.e. the field findings— even if it increases violent behavior relative to exposure to nonviolent movies - i.e. the laboratory findings.

In a related work, Au et al. (2003) estimate that the performance of financial market traders is affected by music-elicited emotional states. The author found that a good mood resulted in inferior performance and overconfidence, while a bad mood resulted in more accurate decisions and more conservative trading. Complimentarily, according to Lerner et al. (2012), sadness increases impatience and creates a myopic focus on obtaining money immediately instead of later. In fact, relative to median neutral-state participants, median sad-state participants accepted 13% to 34% less money immediately to avoid waiting three months for payment.

In another example, Coricelli et al. (2010) finds that the risk of public exposure of deception deters evasion, while the amount of fines encourages evasion. The authors conclude that those results imply that an audit policy that strengthens the emotional dimension of cheating favors compliance (Coricelli et al., 2010). Kushlev et al. (2015),

⁷⁷See for example Troy and Mauss (2011)

using data from a diverse cross section of the U.S. population (N=12,291), shows that higher incomes are associated with experiencing less daily sadness, but have no effect on daily happiness.

A explanation for the asymmetric impact on only the negative reaction is consistent with [Querengsser and Schindler \(2014\)](#)'s findings. In particular, the authors refers to [Nesse \(1990\)](#), who argue that “Emotional states not only motivate action, they are also goals that we seek to achieve. Most human thought, plans, and actions are intended to induce positive emotions or to avoid negative emotions” ([Nesse, 1990](#), p. 262). From this evolutionary point of view, a successful induction of negative emotion would be more relevant for participants' behavior because negative emotions suggest a situation that should be altered, while positive emotions indicate situations that should be maintained ([Nesse, 1990](#); [Querengsser and Schindler, 2014](#)).

Such evidence, along with related lines of work, have contributed to the conclusion that emotion is not epiphenomenal and can influence cognition and behavior in powerful ways.⁷⁸

Finally, the literature has been emphatic arguing that it is crucial to be careful about directly comparing the results of laboratory and field studies ([DellaVigna, 2009](#); [Levitt and List, 2007](#)).

6 Conclusions

This study contributes to the understanding of channels that could be observed —or qualitatively claimed— relative to social programs that foster socio-emotional skills among youth. Consistent with previous studies (e.g. [Attanasio et al., 2015](#); [Calero et al., 2014](#); [Card et al., 2015](#); [West et al., 2015](#)), the Mining's Rockstars participants experience significant impacts on educational outcomes —i.e. dropouts and SAT-like registration rates—, yet no short-run impact among in the expected mechanism, such as perseverance (grit) or internal locus of control. Though evidence of interventions targeted at shaping non-cognitive skills in youth —and thus relieve poverty or improve anti-social behavior— is weak and scarce ([Heckman and Kautz, 2012](#); [Hill et al., 2011](#);

⁷⁸ For reviews, see [Loewenstein et al. \(1992, 2001\)](#); [Damasio \(1994\)](#); [Weber and Johnson \(2009\)](#); [Lempert and Phelps \(2014\)](#); [Lerner et al. \(2015\)](#).

[Blattman et al., 2015](#)), this dissertation shows that social programs designed to foster non-cognitive skills are likely to affect emotional regulation. In other words, emotional regulation can be the mechanism explaining the educational and/or labor market outcomes that are typically observed on assessment of social programs.

I argue that programs to enhance non-cognitive skills likely impact dimensions that are too complex to measure through traditional self-reported psychometric tests. In [Egana-delSol \(2016b\)](#), I claim that emotions are correlated with psychometric test scores typically used to proxy non-cognitive skills. Complementarily, the present study explores the relevance of developing emotional regulation represented by pre-test arousal and valence, as well as emotional response to both positive and negative stimuli. It is worth noting that is hard to disentangle the non-existent impact on non-cognitive skills with the impact on emotions. The program is likely affecting emotion and self-reported tests scores through changes in emotion. In fact, besides the positive qualitative evidence in favor of change of non-cognitive skills, it is not possible to correctly separates these effects. However, there are interesting results.

The neurophysiological measures in the field experiment contribute to two dimensions. First, they make it plausible to hypothesize about the lack of impact of social programs on life skills. This lack of impact is consistent with the program's negative impact on emotional regulation, which is a direct corollary to the positive correlation between emotional state and test scores on non-cognitive tests ([Querengsser and Schindler, 2014](#); [Egana-delSol, 2016b](#)). In other words, the increase in a withdrawal behavior —i.e. less positive valence— biases downward self-reported measures of non-cognitive skills, as showed in [Egana-delSol \(2016b\)](#). In fact, the program might affect the self-perception of subjects at the moment they are answering self-reported tests. There is a behavioral response by students themselves. For instance, a program that encourages grit and locus control in individuals might not change their inherent skills, but may affect their emotional responsiveness to their own environment. It is possible that the treated students become more exigent with themselves than the non treated individuals, in addition to the positive effect that treatment has on their skills.

Second, I find a significant decrease in the neurophysiological reaction to a negative stimulus on the participant, which is a proxy of emotional/behavioral response. There is no significant impact with respect to positive stimuli. These behavioral findings are consistent with the asymmetric impact on emotional responsiveness ([DellaVigna, 2009](#)). For instance, higher income is associated with reduced daily sadness but not

increased daily happiness (Kushlev et al., 2015). I highlight the relevance of emotional disposition and modulation as a key mechanism to overcome "Psychological Poverty Traps" (Haushofer and Fehr, 2014). According to Haushofer and Fehr (2014) poverty causes stress and negative affective states, which may lead to short-sighted and risk-averse decision-making; possibly by, limiting attention and favoring habitual behaviors at the expense of goal-directed ones. This notion of poverty is consistent with recent results in reviews from OECD (2015) and WorldBank (2015).⁷⁹

Moreover, the impacts on emotional responses that I find are likely due to the reappraisal-focused strategies in the program methodologies. Indeed, these strategies have been identified as the most effective way to avoid emotional bias on decision making (Lerner et al., 2015).

The methodology proposed in this study is an alternative to the most frequently used approach in the labor economics literature: latent factor models and the use of revealed behaviors as proxies. The use of latent factor models increases the reliability of skills proxies, but does not solve the problem that the program is potentially affecting other unobservable dimensions such as emotional regulation, and thus self-reported tests. Furthermore, the use of revealed behaviors as proxies, in conjunction with information about past behaviors, is tautological, since it explains behaviors using them as proxies for non-cognitive skills (Heckman and Kautz, 2012).

Moving away from behavioral-based or self-reported measures of non-cognitive skills is a challenging topic for further applications and research. The methodology proposed in this paper has many benefits. First, it offers a way to incorporate emotion into the labor economics field. The importance of emotional regulation —i.e. emotional stability— in labor markets and overall life satisfaction has recently been highlighted for both developed and developing countries (Deming, 2015; OECD, 2015). This study shows that there are neurophysiological approaches to proxy emotional disposition and responsiveness with a high level of accuracy, and at a relatively low cost.

The results also may aid evaluation of similar programs attempting to foster non-cognitive skills. In fact, the effects on emotional regulation —i.e. emotional state and responsiveness— also have implications for the experimental evaluation of educational interventions. Social programs for education and Active Labor Market Policies (ALMP) usually aim to impact socio-emotional or life-skills, such as perseverance, self-control,

⁷⁹ In particular in the reports "Skills for Social Progress: The Power of Social and Emotional Skills" and "The World Development Report: Mind, Society, and Behavior", respectively.

goal-oriented effort, and so forth. However, these factors suffer from measurement error —i.e. reference or emotional bias— due to self-reporting. The methodology proposed here allows us to measure emotional disposition and responsiveness from EEG recordings, which is a non-invasive and low-cost method. Therefore, further research could incorporate physiological measures of emotional state and modulation to study the human capital production function, educational interventions, and the ALMP effectiveness.

References

- Acemoglu, D., 2010. Theory, General Equilibrium, and Political Economy in Development Economics. *Journal of Economic Perspectives* 24 (3), 17–32.
- Albarran, A., Gonzalez, C., May 2015. Introduciendo Habilidades No Cognitivas en Liceos Técnicos Chilenos. Una Evaluación Social. mimeo, Universidad de Chile. Tech. rep.
- Almlund, M., Duckworth, A. L., Heckman, J., Kautz, T., Feb. 2011. Personality Psychology and Economics. NBER Working Paper Series 16822.
- Aspiras, T., Asari, V., 2011. Log power representation of eeg spectral bands for the recognition of emotional states of mind. *Proceedings of the 8th International Conference on Information, Communications and Signal Processing*.
- Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., Rubio-Codina, M., Feb. 2015. Estimating The Production Function For Human Capital: Results From A Randomized Control Trial In Colombia. NBER Working Paper Series 20965.
- Au, K., Chan, F., Wang, D., Vertinsky, I., 2003. Mood in foreign exchange trading: cognitive processes and performance. *Organizational behavior and human decision processes*.
- Banerjee, A., Duflo, E., Nov. 2008. The Experimental Approach To Development Economics. NBER Working Paper Series 14467.
- Barnett, W. S., 2008. Preschool education and its lasting effects: Research and policy implications. Education and Public Interest Center and Education Policy Research Unit.
- Barraza, P., Toro, M., Avaria, R., Oyarzun, F., 2015. Efecto del Contexto Emocional en la Memoria. CIAE Working Paper.
- Bassi, M., Urzua, S., 2010. Educación en Chile: El Desafío está en la Calidad. IADB Technical Notes.
- Bear, M., Connors, B., 2007. *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins.

- Blattman, C., Jamison, J., Sheridan, M., 2015. Reducing adult poverty, crime and violence through late-stage noncognitive investments: Experimental evidence from Liberia. NBER Working Paper Series.
- Borghans, L., Duckworth, A. L., Heckman, J. J., ter Weel, B., 2008. The Economics and Psychology of Personality Traits. *Journal of Human Resources* 43 (4), 972–1059.
- Bos, D. O., 2006. EEG-based emotion recognition. The Influence of Visual and Auditory Stimuli. Unpublished, 1–17.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Routledge & Kegan Paul.
- Brea, M., 2010. Interpretación de las escalas cps, rosenberg y grit y propuestas de revisión para mejorar la confiabilidad. World Bank Report.
- Brown, L., Grundlehner, B., Penders, J., 2011. Towards wireless emotional valence detection from EEG. 33rd Annual International Conference of the IEEE EMBS, 2188–2191.
- Calero, C., Corseuil, C. H., Gonzales, V., Kluge, J., Soares, Y., May 2014. Can Arts-Based Interventions Enhance Labor Market Outcomes among Youth? Evidence from a Randomized Trial in Rio de Janeiro. IZA Discussion Paper Series 8210.
- Cameron, C., Trivedi, P., 2005. *Microeconometrics: methods and applications*. Cambridge University Press.
- Card, D., Ibarrran, P., Regalia, F., Rosas-Shady, D., Soares, Y., Apr. 2011. The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics* 29 (2), 267–300.
- Card, D., Kluge, J., Weber, A., Nov. 2010. Active Labour Market Policy Evaluations: A Meta-Analysis. *The Economic Journal* 120, F452–F477.
- Card, D., Kluge, J., Weber, A., Jul. 2015. What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. IZA Discussion Paper Series 9236.
- Carrel, S., Sacerdote, B., 2013. Why do college going interventions work? NBER Working Paper Series (19031).

- Castellanos, N. P., Makaro, V. A., 2006. Recovering eeg brain signals: Artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods* 158, 300–312.
- Choppin, A., 2000. EEG-Based Human Interface for Disabled Individuals: Emotion Expression with Neural Networks. Unpublished Master Thesis.
- Coombes, S., Cauraugh, J., Janelle, C., 2007. Dissociating Motivational Direction and Affective Valence Specific Emotions Alter Central Motor Processes. *Psychol Sci* 18 (11), 938–942.
- Coricelli, G., Joffily, M., Montmarquette, C., Villeval, M. C., 2010. Cheating, Emotions, and Rationality: An Experiment on Tax Evasion. *Experimental Economics* 13 (2), 226–247.
- Cox, C., 2006. Policy Formation And Implementation in Secondary Education Reform: The Case of Chile at the Turn of the Century. *The Education Working Paper Series* 3.
- Cullen, J. B., Jacob, B. A., Levitt, S., Sep. 2006. The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica* 74 (5).
- Cunha, F., Heckman, J., Schennach, S., May 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica* 78, 883–931.
- Cunha, F., Heckman, J. J., Jan. 2007. The Technology of Skill Formation. *IZA Discussion Paper Series* 2550.
- Cunha, F., Heckman, J. J., 2008. **Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation** . *Journal of Human Resources* 43 (4), 738–782.
- Dahl, G., DellaVigna, S., 2009. Does Movie Violence Increase Violent Crime? *Q J Econ* 124, 677–734.
- Damasio, A., 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam Publishing.
- Darwin, C., 1872. *The Expression of the Emotions in Man and Animals*. John Murray.

- Davidson, R., Schaffer, C., Saron, C., Jul. 1983. Frontal and Parietal Electroencephalogram Asymmetry in Depressed and Nondepressed Subjects. *Biol Psychiatry* 18 (7), 753–762.
- Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., Friesen, W. V., 1990. Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of personality and social psychology* 58 (2), 330.
- Deaton, A., 2010. Instruments, Randomization, and Learning About Development. *Journal of Economic Literature* 48, 424–455.
- DellaVigna, S., 2009. Psychology and Economics: Evidence from the Field. *Journal of Economic Literature* 47 (2), 315–372.
- Delorme, A., Makeig, S., 2004. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics. *Journal of Neuroscience Methods* 134, 9–21.
- Deming, D. J., Aug. 2015. The Growing Importance of Social Skills in the Labor Market. NBER Working Paper Series 21473.
- Descartes, R., 1649. *The Passion of the Soul*. Hackett Publishing Company.
- Duckworth, A. L., Peterson, C., Matthews, M. D., Kelly, D. R., 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology* 92 (6), 1087–1101.
- Duflo, E., Glenner, R., Iqbal, D., 2012. Comparative cost- effectiveness analysis to inform policy in developing countries: A general framework with application. MIT JPAL Policy Working Papers.
- Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., Schellinger, K., 2011. The impact of enhancing students’ social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development*.
- Edmans, A., Garcia, D., Norli, Ø., 2007. Sports Sentiment and Stock Returns. *The Journal of Finance* 62 (4).
- Egana-delSol, P., 2016a. Can art-based programs nurture human capital among youth? evidence from a quasi-experimental trial. PhD Thesis Chapter, Columbia University.

- Egana-delSol, P., 2016b. Can low-cost electroencephalogram (eeg) devices measure human capital characteristics more effectively? PhD Thesis Chapter, Columbia University.
- Ekman, P., Friesen, W., O'Sullivan, M., Diacoyanni-Tarlatzis, I., Krause, R., Pitcairn, T., Scherer, K., Chan, A., Heider, K., LeCompte, W., Ricci-Bitti, P., Tomita, M., Tzavaras, A., Darwin, C., 1987. Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion. *Journal of Personality and Social Psychology* 53 (4), 712–717.
- Fazio, M., 2011. Análisis de la percepción de los empleadores acerca de las pasantías y las perspectivas de inserción laboral en esas empresas. Technical Note, InterAmerican Development Bank IDB-TN-240.
- Friedman, D., Shapira, S., Jacobson, L., Gruberger, M., 2015. A data-driven validation of frontal eeg asymmetry using a consumer device. 2015 International Conference on Affective Computing and Intelligent Interaction (ACII).
- Fuster, J., 2002. Frontal Lobe and Cognitive Development. *Journal of Neurocytology*, 373–385.
- Fuster, J., 2013. *The Neuroscience of Freedom and Creativity*. Cambridge University Pres.
- Gintis, H., 1971. Education, technology, and the characteristics of worker productivity. *American Economic Review* 61 (2).
- Goleman, D., 2010. *Emotional intelligence: Why it can matter more than iq*. Bantam Books.
- Gross, J. J., John, O. P., Kaufman, J., 2003. Individual Differences In Two Emotion Regulation Processes: Implications For Affect, Relationships, And Well-Being. *Journal of Personality and Social Psychology* 85 (2), 348–362.
- Harmon-Jones, E., Gable, P. A., 2008. Approach-Motivated Positive Affect Reduces Breadth of Attention. *Psychol Sci* 19 (5), 476–482.
- Harmon-Jones, E., Gable, P. A., Peterson, C. K., 2010. The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biol Psychol* 84 (3), 451–462.

- Haushofer, J., Fehr, E., 2014. On the Psychology of Poverty. *Science* 344 (186), 862–863.
- Heckman, J. J., Kautz, T., 2012. Hard Evidence on Soft Skills. *Labour Econ* 19 (4), 451–464.
- Heckman, J. J., Stixrud, J., Urzua, S., 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. NBER Working Paper Series 12006.
- Hill, P. L., Roberts, B. W., Grogger, J. T., Guryan, J., Sixkiller, K., Jan. 2011. Decreasing Delinquency, Criminal Behavior, and Recidivism by Intervening on Psychological Factors Other than Cognitive Ability: A Review of the Intervention Literature. NBER Working Paper Series 19656.
- Hirschi, T., Gottfredson, M., Feb. 1993. Commentary: Testing the General Theory of Crime. *Journal of Research in Crime and Delinquency* 30 (1), 47–54.
- Hughes, D., Yates, M., Morton, E., Smillie, L., Dec. 2014. Asymmetric frontal cortical activity predicts effort expenditure for reward. *Soc Cogn Affect Neurosci* 10 (7).
- Ibarraran, P., Ripani, L., Taboada, B., Villa, J. M., Garcia, B., Jun. 2014. Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design. *IZA Journal of Labor Development* 3 (10), 1–24.
- Jackson, C. K., 2014. Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. Untitled. NBER Working Paper Series.
- Jamieson, J. P., Nock, M. K., Mendes, W. B., 2012. Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology* 141 (3), 417–422.
- John, O., Srivastava, S., 1999. Big Five Inventory (BFI). *Handbook of personality: Theory and research*. Pervin, L A and John, O P Editors. 1, 102–138.
- Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G., Just, M. A., Jun. 2013. Identifying Emotions on the Basis of Neural Activation. *PLoS One* 8 (6), e66032.
- Kaufman, J., Sternberg, R., 2010. *The Cambridge Handbook of Creativity*. Cambridge University Press.

- Kautz, T., Heckman, J., Diris, R., Borghans, L., ter Weel, B., 2014. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. OECD Education Working Papers (110).
- Keenan, S. A., 1999. Normal human sleep. *Respiratory Care Clinics of North America* 5 (3), 319–31, vii.
- Kim, M., Kim, M., Oh, E., Kim, S., 2013. A Review on the Computational Methods for Emotional State Estimation from the Human EEG. *Comput Math Methods Med* 2013, 1–13.
- Klimesch, W., 1996. Memory Processes, Brain Oscillations And EEG Synchronization. *International Journal of Psychophysiology* 24, 61–100.
- Klimesch, W., 1999. Klimesch W. Eeg Alpha And Theta Oscillations Reflect Cognitive And Memory Performance: A Review And Analysis. *Brain Research Reviews* 29 (2), 169–195.
- Kluve, J., 2010. The effectiveness of European active labor market programs. *Labour Econ* 17 (6), 904.
- Knyazev, G. G., 2007. Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neurosci Biobehav Rev* 31 (3), 377–395.
- Koelsch, S., Apr. 2012. *Brain and Music*. Wiley-Blackwell.
- Korjus, K., Uusberg, A., Uusberg, H., Kuldkepp, N., Kreegipuu, K., Allik, J., Vicente, R., Aru, J., Feb. 2015. Personality Cannot Be Predicted from the Power of Resting State EEG. *Front Hum Neurosci* 9 (63), 1–7.
- Kushlev, K., Dunn, E., Lucas, R., Jul. 2015. Higher Income Is Associated With Less Daily Sadness But Not More Daily Happiness. *Social Psychological and Personality Science* 6 (5), 483–489.
- Lakoff, G., 2008. *The Political Mind: Why You Can't Understand 21st-Century Politics With An 18th-Century Brain*. Viking.
- Lempert, K., Phelps, E., 2014. Neuroeconomics of emotion and decision making. In: Glimcher, P., Fehr, E. (Eds.), *Neuroeconomics: Decision Making and the Brain*, 2nd Edition. Academic Press, San Diego, USA., Ch. 12, pp. 219–236.

- Lerner, J., Li, Y., Weber, E., 2012. The Financial Costs of Sadness. *Psychol Sci* 24 (1), 72–79.
- Lerner, J. S., Li, Y., Valdesolo, P., Kassam, K. S., Madrian, B., Shea, D., 2015. Emotion and Decision Making. *Annual Review of Psychology* 66 (1), 799–823.
- Levitin, D., 2006. *This Is Your Brain on Music*. Dutton Penguin.
- Levitt, S., List, J., 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives* 21 (2), 153–174.
- Lin, Y., Wang, C., Jung, T., Wu, T.-L., Jeng, S., Duann, J., Chen, J., 2010. EEG-Based Emotion Recognition in Music Listening. *IEEE Transactions on Biomedical Engineering* 57 (7), 1798–1806.
- Loewenstein, G., 2000. Emotions in economic theory and economic behavior. *The American Economic Review* 90 (2).
- Loewenstein, G., Prelec, D., Loewenstein, G., Prelec, D., May 1992. Anomalies in Intertemporal Choice: Evidence and an Interpretation. *Q J Econ* 107 (2), 573–597.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., Welch, N., 2001. Risk as feelings. *Psychological Bulletin* 127 (2), 267–286.
- Martinez-Leon, J.-A., Cano-Izquierdo, J.-M., Ibarrola, J., 2016. Are low cost brain computer interface headsets ready for motor imagery applications? *Expert Systems with Applications* 49 (1), 136–144.
- McClelland, D., Jan. 1973. Testing for Competence Rather Than for "Intelligence". *American Psychologist* 8 (1), 1.
- McKenzie, D., Feb. 2014. Hard measurement of Soft Skills. *World Bank Development Blog*.
- McKenzie, D., Groh, M., Vishwanath, T., 2014. Reducing information asymmetries in the youth labor market of Jordan with psychometrics and skill based tests. *ABCDE Conference*.
- Mitra, P., Bokil, H., 2008. *Observed Brain Dynamics*. Oxford University Press, New York.

- Moretti, D., Babiloni, C., Binetti, G., Casetta, E., Dal Forno, G., Ferreric, F., Ferri, R., Lanuzza, B., Miniussi, C., Nobili, F., Rodriguez, G., Salinari, S., Rossini, P., 2004. Individual Analysis Of Eeg Frequency And Band Power In Mild Alzheimer'S Disease. *Clinical Neurophysiology* 115 (2), 299–308.
- Morrison, L., Shoon, I., 2013. The impact of non-cognitive skills on outcomes for young people . Institute of Education, University of London. The Education Endowment Foundation.
- Murnane, R., Ganimian, A., Jul. 2014. Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations. NBER Working Paper Series 20284.
- Nesse, R. M., 1990. Evolutionary explanations of emotions. *Human Nature* 1 (3), 261–289.
- Ochsner, K., Bunge, S., Gross, J., Gabrieli, J., 2002. Rethinking feelings: an fmri study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience* 14 (1).
- OECD, 2015. Skills for Social Progress: The Power of Social and Emotional Skills. OECD Publishing, OECD Skills Studies.
- ONeil, H., Perez, R., Baker, E., 2014. Teaching and Measuring Cognitive Readiness. Springer.
- Partala, T., Jokiniemi, M., Surakka, V., Jan. 2000. Pupillary responses to emotionally provocative stimuli. *Eye Tracking Research and Applications Symposium Proceedings*.
- Petrantonakis, P., Hadjileontiadis, L., Mar. 2010. Emotion Recognition From EEG Using Higher Order Crossings. *IEEE Transactions on Information Technology in Biomedicine* 14 (2), 186–197.
- Pierre, G., Sanchez-Puerta, M., Valerio, A., Rajadel, T., 2014. STEP Skills Measurement Surveys: Innovative Tools for Assessing Skills. *Social Protection and Labor Discussion Papers*.
- Plucker, J., Makel, M., 2010. Assessment of Creativity. In: Kaufman, J., Sternberg, R. (Eds.), *The Cambridge Handbook of Creativity*. Cambridge University Press, Cambridge, United Kingdom, pp. 48–73.

- Pop-Eleches, C., Urquiola, M., 2013. Going to a Better School: Effects and Behavioral Responses. *American Economic Review* 103 (4), 1289–1324.
- Pratt, T., Cullen, F., 2000. The Empirical Status of Gottfredson and Hirschi's General Theory of Crime: A Meta-Analysis. *Criminology* 38 (2).
- Puentes, E., Urzua, S., Dec. 2010. La Evidencia del Impacto de los Programas de Capacitación en el Desempeño en el Mercado Laboral. IADB Technical Notes.
- Querengsser, J., Schindler, S., 2014. Sad but true? - How induced emotional states differentially bias self-rated Big Five personality traits. *BMC Psychology* 2 (4).
- Ramirez, R., Vamvakousis, Z., 2012. Detecting emotion from EEG signals using the emotive epc device. *Brain Informatics*, 175–184.
- Raven, J., 1936. Mental Tests Used in Genetic Studies: The Performance of Related Individuals on Tests Mainly Educative and Mainly Reproductive. Master Thesis.
- Rotter, J. B., 1966. Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs: General and Applied* 80 (1), 1–28.
- Runco, M., 2010. Divergent thinking, creativity, and ideation. In: Kaufman, J., Sternberg, R. (Eds.), *The Cambridge Handbook of Creativity*. Cambridge University Press, Cambridge, United Kingdom, Ch. 22, pp. 413–446.
- Salzman, C. D., Fusi, S., 2010. Emotion, Cognition, and Mental State Representation in Amygdala and Prefrontal Cortex. *The Annual Review of Neuroscience* 33, 173–202.
- Sigman, M., Peña, M., Goldin, A. P., Ribeiro, S., Mar. 2014. Neuroscience and education: prime time to build the bridge. *Nature Science Review* 17 (4), 497–502.
- Sobolewski, A., Holt, E., Kublik, E., Wróbel, A., 2011. Impact of meditation on emotional processing: A visual ERP study. *Neuroscience Research* 71 (1), 44–48.
- Sokol-Hessner, P., Camerer, C., Phelps, E., Feb. 2012. Emotion Regulation Reduces Loss Aversion and Decreases Amygdala Responses to Losses. *Soc Cogn Affect Neurosci* 8 (3), 341–350.
- Spielberg, J. M., Stewart, J. L., Levin, R. L., Miller, G. A., Heller, W., 2008. Prefrontal Cortex, Emotion, and Approach/Withdrawal Motivation. *Social and Personality Psychology Compass* 2 (1), 135–153.

- Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., Okubo, Y., 2004. Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *Neuroimage* 23, 967–974.
- Todd, P., Wolpin, K., 2003. On The Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal* 113, F3–F33.
- Torrance, E. P., 1966. *Torrance Tests of Creative Thinking: Norms technical manual (Research Edition)*. Personnel Press.
- Troy, A. S., Mauss, I. B., 2011. Resilience in the face of stress: emotion regulation as a protective factor. In: Southwick, S. M., et al. (Eds.), *Resilience and Mental Health: Challenges Across the Lifespan*. Cambridge University Press.
- Valenzuela, J. P., Allende, C., Sevilla, A., Egana del Sol, P., May 2013. La (ina)movilidad del desempeño educativo de los estudiantes chilenos: realidad, oportunidades y desafíos. Documento de Trabajo CIAE.
- Verma, G., Tiwary, U. S., 2014. Multimodal Fusion Framework: A Multiresolution Approach for Emotion Classification and Recognition from Physiological Signals. *Neuroimage* 102, 162–172.
- Wacker, J., Heldmann, M., Stemmler, G., 2003. Separating Emotion and Motivational Direction in Fear and Anger: Effects on Frontal Asymmetry. *Emotion* 3, 167–193.
- Watson, D., Clark, L. A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology* 54 (6), 1063–1070.
- Weber, E. U., Johnson, E. J., 2009. Mindful judgment and decision making. *Annual Review of Psychology* 60, 53–85.
- West, M., Kraft, M., Finn, A., Martin, R., Duckworth, A., Gabrieli, C., Gabrieli, J., 2015. Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*.
- West, M. R., Kraft, M. A., Finn, A. S., Duckworth, A. L., Gabrieli, C. F., Gabrieli, J. D. E., 2014. Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. Center for Education Policy Research WP, Harvard University.

WorldBank, 2015. The World Development Report: Mind, Society, and Behaviour. World Bank.

Yoon, H. J., Chung, S. Y., 2013. EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm. *Computers in Biology and Medicine* 43, 2230–2237.

7 Appendix 1: Balance of Attrition

Table 8: Balance of Attrition (missed subjects compared at baseline)

Orthogonality Table				
	Control	Treated	(1) vs. (2)	p-value
Locus Control Test	-0.016 (0.113)	0.270 (0.117)	-0.286 (0.167)	0.089
Grit Test	0.053 (0.115)	0.312 (0.110)	-0.260 (0.165)	0.119
Creativity Index	0.111 (0.135)	-0.002 (0.162)	0.113 (0.211)	0.594
13 Forces Creativity	-0.001 (0.131)	-0.095 (0.152)	0.094 (0.202)	0.641
Standardized values of (valen_neg)	-0.016 (0.160)	0.132 (0.131)	-0.148 (0.221)	0.506
Standardized values of (valen_pos)	-0.031 (0.135)	0.053 (0.114)	-0.084 (0.188)	0.657
Standardized values of (valen_base)	0.189 (0.109)	0.061 (0.091)	0.128 (0.152)	0.401
Standardized values of (arousal_base)	-0.088 (0.127)	0.361 (0.177)	-0.449 (0.212)	0.036
N	66	46	112	

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

8 Appendix 2: Robustness Checks

8.1 Robustness Check: Impact on Cognitive Skills (Raven-like Test)

Table 9 presents the program’s impact on cognitive skills based on the Raven’s progressive matrices test, which capture logical and problem/solving skills through guessing the missing matrix between many Raven’s-like progressive matrices. The Mining’s Rockstars program is not designed to affect cognitive skills, thus the test was incorporated only in the follow up survey as a placebo test. Regardless, since the program is affecting emotional state and responsiveness, we could expect an indirect effect, but as Table 9 shows, this is not the case here either.

Table 9: Placebo Test - Raven-like Test and Grade Previous Year

VARIABLES	(1) Raven Test	(2) Grade Avg.
Treat & Post.	-0.0530 (0.0525)	0.193 (0.153)
Treatment		-0.127 (0.154)
Post.		-0.108 (0.111)
Constant	0.554*** (0.0705)	5.805*** (0.102)
Observations	116	166
R-squared	0.102	0.196
School Dummies	Yes	Yes
Schools's Class Cluster	Yes	Yes

Robust p-values in parentheses. ***

p<0.01, ** p<0.05, * p<0.1.

Notes: "Grade Avg." is a variable that accounts for the average grade obtained the year before the baseline was conducted.

8.2 Robustness Check: In Difference Model: Impact on Emotional Regulation

Table 10 exposes the program's impact following a Model in differences (Cameron and Trivedi, 2005). The model considers two novel outcomes attempting to capture the unobservable impact on emotional regulation on a subsample that considers only those individuals that did both baseline and follow up field experiments— i.e. the balanced panel. This model allows to clean all time-invariant non-observable characteristics on both participant and their control group. In particular, the following model is estimated:

$$\Delta E_{j,i} = \delta_1 * treat_i * Post_t + \theta_1 * X_i + \varepsilon_{1i,t} \quad (14)$$

where $E_{j,i,t}$ indicates the j emotional state or emotional responsiveness associated with an emotionally laden stimuli of individual i in time t , with $j = arousal_{base}$,

$valence_{base}$, $valence_{positive}$, $valence_{neg}$, $dif-valence_{pos}$, $dif-valence_{neg}$, $i = students \in FieldExperiment$, and $t = baseline, follow-up$. As before, $treat_{j,i}$ and $Post_{j,t}$ indices who had been treated by the program and survey time —i.e. baseline or follow up— respectively. Finally, $X_{i,t}$ indicates school level dummy variables. It is expected that $\delta_1 < 0$ and $\beta = \gamma = 0$.

8.3 Robustness Check: Wild Bootstrap on p-values (Exact p-values)

Table 11: Robustness Check: Wild Bootstrap on Emotions

VARIABLES	(1) ArouRest	(2) ValeRest	(3) ValNeg	(4) ValPos	(5) ValNegDif	(6) ValPosDif
Treat & Post.	-0.532** (0.0224)	-1.408* (0.0940)	0.724 (0.281)	-0.0274 (0.974)	-2.131** (0.0353)	-1.380 (0.250)
Treatment	0.0588 (0.766)	0.107 (0.811)	0.00262 (0.997)	0.597 (0.325)	0.105 (0.883)	-0.489 (0.501)
Post.	-0.00964 (0.946)	0.273 (0.675)	-1.131** (0.0251)	0.476 (0.426)	1.404* (0.0727)	-0.203 (0.811)
Constant	-0.0848 (0.663)	0.173 (0.775)	-0.263 (0.752)	-1.418 (0.101)	0.436 (0.634)	1.591* (0.0999)
Observations	296	296	296	296	296	296
R-squared	0.050	0.023	0.037	0.017	0.033	0.027
School Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Student Level Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Exact pvalue (WildB.)	0.0180	0.0860	0.270	0.971	0.0360	0.243

Robust p-values in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: Exact p-value is the results of 1,000 replications of Wild bootstrap, following the algorithm of Webb for few clusters. Results are similar if clusterization is at school level.

Table 10: Impact on Emotional Regulation: Emotional State and Responsiveness

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ArouRest	ValRest	ValNeg	ValPos	ValNegDif.	ValPosDif.	Delta_ValPosDif.	Delta_ValNegDi
Treat & Post.	-0.591** (0.233)	-1.350* (0.802)	0.649 (0.636)	-0.0524 (0.802)	-1.999** (0.954)	-1.298 (1.135)	-2.932 (2.679)	-3.383* (2.011)
Treatment	0.0669 (0.196)	0.129 (0.436)	0.0130 (0.592)	0.597 (0.571)	0.116 (0.669)	-0.468 (0.695)		
Post.	0.00761 (0.155)	0.255 (0.626)	-1.079** (0.470)	0.468 (0.568)	1.335* (0.736)	-0.213 (0.804)		
Constant	-0.0639 (0.190)	0.191 (0.588)	-0.219 (0.774)	-1.319 (0.814)	0.410 (0.857)	1.510 (0.924)	-0.831 (2.282)	3.493 (2.798)
Observations	296	296	296	296	296	296	68	68
R-squared	0.059	0.024	0.038	0.016	0.034	0.027	0.102	0.166
Student FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Student Clust.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Robustness Check: Wild Bootstrap on Non-Cognitive Skills

VARIABLES	(1) Creativity	(2) 13Forces Creat.	(3) LocusControl Test	(4) Grit Test
Treat & Post.	0.243 (0.340)	0.172 (0.466)	-0.105 (0.668)	0.0994 (0.626)
Treatment	0.0300 (0.880)	0.0424 (0.833)	0.327 (0.214)	0.706*** (0.00153)
Post.	0.0545 (0.745)	0.318* (0.0638)	0.166 (0.307)	-0.0179 (0.903)
Constant	-0.0811 (0.715)	-0.179 (0.425)	0.0129 (0.947)	-0.183 (0.270)
Observations	296	296	296	296
R-squared	0.074	0.086	0.029	0.110
Student FE	Yes	Yes	Yes	Yes
Student level Cluster	Yes	Yes	Yes	Yes
School Dummies	Yes	Yes	Yes	Yes
Exact pvalue (WildB.)	0.333	0.477	0.667	0.624

Robust p-values in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes: Exact p-value is the results of 1,000 replications of Wild bootstrap, following the algorithm of Webb for few clusters. Results are similar if clusterization is at school level.

9 Appendix 3: Field Experiment Setting

Figure 5: Field Experiment Setting

