Article

Not peer-reviewed version

The Combined Effect of Big 5 Personality Traits on Fourth-Graders´ Math Performance

Roberto Araya * and Pablo González-Vicente

Posted Date: 18 August 2023

doi: 10.20944/preprints202308.1302.v1

Keywords: Big 5; Child Personality; Elementary School Mathematic Performance; Socio Emotional Effects; Educational Data Mining

Article

The Combined Effect of Big 5 Personality Traits on Fourth-Graders´ Math Performance

Roberto Araya [1,*] and Pablo González-Vicente [2]

[1] Institute of Education; roberto.araya.schulz@gmail.com
[2] Institute of Education; pgonzalezvicente1@gmail.com
* Correspondence: roberto.araya.schulz@gmail.com; Tel.: +56 99 599 0251

Abstract: The "Big Five" model of personality has been widely validated in adults, but its application to children is challenging. An important advance is the 15-questions Pictorial Personality Traits Questionnaire for Children. A study on students from 10 to 13 years old in Poland validated the questionnaire but with some observations. Thus, there is a need for replication and stronger evidence in this age group. In Chile, we replicated the study with 3,423 4th-graders (9 to 12 years old). Teachers, in regular sessions, applied the questionnaire to their entire classes. We found similar results, including that asexual pictograms worked well in both genders. We also found positive relationships between conscientiousness, openness, and extraversion with mathematical performance. Furthermore, a combination of these three traits has a relationship with math performance twice as big as each trait alone. Moreover, students with the lowest scores in this combination of personality traits (6.6% of students) have 0.27 standard deviations less in mathematical performance than those with the highest score, which is 74.3% of students. To the best of our knowledge, this is the first time that a study finds a strong relationship between a combination of personality traits gathered with a 15-question questionnaire and fourth-graders´ math performances.

Keywords: Big 5; child personality; elementary school mathematic performance; socio emotional effects; educational data mining

1. Introduction

The study of personality has been a central theme of psychology since its origins. There are various approaches, but the psychometric theory of personality traits has developed enormously in the last century, with the pioneering work of Raymond Cattel, Hans J. Eysenck, Robert Plomin, and many others [1]. These works make use of advanced mathematical and statistical techniques, particularly factor analysis, for which its popularization had to wait for the advent of computers and adequate software developments to implement analyzes in a bigger scale.

Thanks to numerous empirical studies in adults, there is now a broad consensus that there are five personality traits [2]. These traits are agreeableness, conscientiousness, extraversion, neuroticism, and openness to experience. The Big 5 model with these five personality traits is pretty universal. Several studies has validated the model in different cultures and some of its components in other animals such as chimpanzees [3].

However, in some developing countries (Botswana, Ethiopia, Lebanon, Malaysia, Puerto Rico, and Uganda), the factor structure was not so evident [3]. [3] argues that one possible explanation is that in these countries, the quality of the data was quite poor, people may have misunderstood the questions or may not be familiar with the question format.

In addition, using this model in children has presented great challenges, since the measurement instruments applied in adults are not suitable for preadolescent children [4]. The study of this model in children of different ages is key to help understand how personality traits appear and differ as the individual matures [5]. In contrast, in adulthood, these 5 traits are stable.

 doi:10.20944/preprints202308.1302.v1

In recent years, research on children has expanded, developing special measurement instruments for children and adolescents, but always maintaining its abstract approach based on verbal formulations. For example, the BFQ-C questionnaire (Big Five Personality Questionnaire for Children) consists of 65 items arranged on a five-point Likert-type scale, with values between one point "Completely false for me" and five points "Completely true to me." [6]. This questionnaire is available in several languages. For example, it was adapted to Spanish by [7]. Since then, several studies in Spain and other Spanish-speaking countries have used this translation [8]. However, this approach has received criticism because in children under 12 years of age, verbal and reading-writing skills are not yet fully developed [4].

Another approach has been to have the questionnaire answered by adults (parents or teachers, for example) on behalf of the children. This approach has also received criticism, since it is not sure that adults represent children with high fidelity.

An important innovation is the development of the Pictorial Personality Traits Questionnaire for Children (PPTQ-C), proposed by [4]. This questionnaire uses pictorial scales, with pictographs aligned to the level of cognitive development of children. In addition, the number of items considered for each personality trait has been considerably reduced, going from 13 items in the adult questionnaire to only 3 items in the children's questionnaire.

[4] carried out a study in 527 students from 10 to 13 years old in Poland, validating the questionnaire and the “Big Five” model in this group of children, but with certain apprehensions. In particular, they conclude that "openness to experience" presents weaker validation. Therefore, there is a need for stronger evidence in order to validate the “Big Five” model in children, repeating the study in different countries and with larger samples.

This replication is particularly important, given the growing concern in psychology of a replicability crisis. In a 2016 online survey answered by 1,576 researchers, more than 70% have tried and failed to reproduce another researcher´s experiment, and more than 50% failed to reproduce one of their own experiments [9]. [10] found that only 23% of the replicas achieve the results published in the original articles of a reputed social psychology journal. The probability that a research claim is true depends on the number of subjects, potential biases, effect sizes, and various other factors [11, 12]. For this reason, it is very important to repeat the study with the 15-question PPTQ-C questionnaire, to do it in another population of children and to carry it out in a larger sample. There is also a need to assess whether to improve the PPTQ-C questionnaire and verify if the 5 personality traits are well differentiated between children.

In addition, we are interested in studying whether personality factors have some relationship with the academic performance of students. There is some evidence that trait-based personality types in children, like resiliency and control, are predictors of future outcomes in both the social and cognitive domains [13]. However, for public policies in education, it is interesting to know the effect on specific subject matters, critical in the national curriculum. In particular, it is important the case of mathematics performance for two reasons. First, this is one of the two central subjects that all countries measure. For example, international organizations like OECD and UNESCO, periodically measure math performances in several countries, compare results and track the trends throughout years and decades. Second, there is abundant evidence of a trend in increase of mood disorders in adolescents, like anxiety and depression [14], and particularly, that mathematics education generates emotional disorders in part of the student population [15, 16]. Therefore, it is very important to find out whether there are clusters of students with specific personality profiles that can most affect their performance in mathematics. There are several potential explanations for the increasing appearance of emotional disorders due to mathematics education [17, 18]. Thus, it is of particular interest to investigate the relationship between personality profiles and performance in mathematics.

In summary, our research questions are:

Q1: Is it feasible to do this measurement in large scale during regular classroom sessions? That is, it is possible that a teacher apply the questionnaire, on an online platform, on a regular session, and with a previous virtual training of just 20 minutes.

 doi:10.20944/preprints202308.1302.v1

Q2: Is the 15-questions pictorial scales questionnaire reliable for the study of child personality using the "Big Five" model?

Q3: Is there a relationship between the big 5 personality traits and performance in elementary school mathematics? Moreover, if there is, how does each trait influence and to what extent?

2. Related work

The "Big Five" model of personality traits is one of the most widely used and has been the subject of an increasing number of investigations, especially to study its relationship with cognitive abilities and school-academic success, its relationship with social skills and its ability to predict antisocial behaviors.

Towards the end of 2019, OECD carried out a study in 10 cities in 9 countries around the world, where 3,000 10-year-old students and 3,000 15-year-old students were surveyed [19]. OECD administered an online survey to students using digital devices. They also administered surveys to the parents and teachers of the sampled students, in order to complement and contextualize the information provided by the students. This OECD Survey on Social and Emotional Skills (SSES) is one of the first and most ambitious international efforts to collect data from students, parents and teachers on students' social and emotional skills, in order to guide public policies. Pre-survey work included extensive background information gathering and analysis of the literature regarding social and emotional skills, resulting in the *Social and Emotional Skills for Student Success and Well Being: Conceptual Framework for the OECD Study on Social and Emotional Skills (SSES)* report published by the OECD in 2018 [20]. This report argues that the Big Five model not only provides a comprehensive and accurate summary of individual differences in social and emotional skills, but also has a strong empirical basis, high predictive power, and high inter-cultural stability. For this reason, they used it as a frame of reference for the SSES study carried out the following year. In addition, they operationalized the 5 personality traits in specific social and emotional skills.

The OECD SSES study is interesting not only because it includes measures of socio-economics, school performance, psychological well-being, bullying and school adaptability, but also includes a wide range of countries: Finland, Portugal, USA, Canada, Turkey, the Russian Federation, South Korea, the People's Republic of China and Colombia (with 2 cities). This study clearly and forcefully confirms that social and emotional skills associated with personality traits are a strong predictor of school grades. It also found that those skills are the most significant correlate of students' future educational expectations. Likewise, it shows that there is a strong relation between socioemotional skills of the students and their psychological well-being, discounting the effects of the socioeconomic situation and gender. They detected notable gender differences, similar to other studies. On average, boys reported higher emotional regulation, sociability, and energy levels, while girls reported higher levels of responsibility, empathy, and cooperation with others. The report concludes that emotional and social skills influence a broad set of people's life outcomes and states "the OECD's long-standing commitment to support countries in their efforts to reorient their educational policy agendas to focus more on social and emotional skills and, by extension, develop well-rounded citizens."

A recently published study [21] confirms and sheds new light on the correlation between personality traits and cognitive abilities. In this case, it is a large-scale meta-analysis, based on 1,976 independent samples (from more than 50 countries) and 60,690 correlations between personality traits and cognitive ability. This implies data from millions of individuals, where the age of the participants ranged from 12 to 100 years of both genders (54.1% male). The interesting thing about this enormous work is that it uses the Big Five hierarchical model, distinguishing 2 sublevels (aspects) for each personality trait and specific facets for each aspect; the same for cognitive abilities, where 4 areas or domains are distinguished, each of them broken down into specific skills. Therefore, the richness of the study comes from the analysis of the cross-correlations between specific facets of personality and specific cognitive abilities. These specific correlations have been less studied. This study finds that cognitive abilities are not only affected by the personality traits of open-mindedness and conscientiousness, as was already known, but that there is a broad and very rich network of

 doi:10.20944/preprints202308.1302.v1

subtle relationships between the different facets of personality and the various specific abilities that constitute cognitive ability.

[22] in a recent meta-analysis, studies the relationship between the 5 personality factors of the Big Five model and school performance in particular. This meta-analysis considers 78 studies, 110 samples of students from elementary to high school, representing data from more than 500,000 individuals. 1,491 effect sizes were analyzed using a random effects model and a robust estimate of variance to estimate the mean sizes of each effect. The authors want to delve into the relationships between personality traits and academic performance, taking into account, different subject domains (for example STEM versus language) and the different measures of achievement (for example standardized tests versus grades). The results of the meta-analysis conclude that there are indeed differences worth taking into consideration. For example, the impact of the openness trait is more relevant for academic achievement in the domains of linguistic subjects than for STEM, and this result is valid both when measuring performance with standardized tests and with GPAs placed by each teacher. Instead, the impact of conscientiousness is stronger with grades than with standardized tests, but the effect size is similar across subject areas or subjects.

Most of the research are on adults and college students. In some of them adolescents have been included. However, in children there is still very little evidence, despite the fact that this type of research is essential to guide educational policies in the early stages of schooling. Personality traits evolve with age, as many studies have confirmed. In a meta-analysis carried out in 2006, based on 92 longitudinal studies, [23] reported such changes. Some personality traits grow into adulthood. This is the case with conscientiousness, extraversion and emotional stability, which grow mainly in adulthood (20 to 40 years, after having suffered a sharp decline in adolescence). On the other hand, measures of social vitality (a facet of extraversion) and openness to experience grow strongly in adolescence, but then decrease towards old age.

Studies in preadolescent children are much scarcer. In 1999 in the Netherlands [24] carried out one of the first works to validate the Big Five model in children. 367 children from 4 to 12 years old, chosen from 46 schools, were rated by their teachers, using the scales of the revised "School Behavior Checklist" (SCHOBL-R1). The analysis produced four stable significant components for age and sex: Extraversion, Attitude towards School Work, Agreeableness and Emotional Stability. Despite the fact that the questionnaire items use other personality scales, the authors concluded it is possible to interpret the four factors found with SCHOBL-R in terms of four of the five personality dimensions of the Big Five Model. This finding supports the idea that Big 5 is plausible in Dutch schoolchildren of these ages [24].

It is worth to note that the OECD study carried out in 2019, which includes 10-year-olds (and another group of 15-year-olds), used the Big Five model as a frame of reference. OECD administered an online questionnaire and lasted an average of 46 minutes. The questionnaire is essentially of the same type that the one used with adults. That is, it uses standard metrics and uses a linguistic support that assumes developed reading and writing skills.

On the other hand, the large study carried out in China in 2021 on students between the ages of 6 and 12, contemplated 10,366 students chosen from 21 primary schools in the north of that country. In this case, the authors used the "Personality Inventory for Elementary School Students" [25], which includes five dimensions: extraversion, agreeableness, conscientiousness, emotional stability, and intelligence (the Chinese equivalent of openness). This inventory includes 62 items, rated on a five-point Likert scale from 1 (very inaccurate) to 5 (very accurate). However, the personality traits of each student were evaluated by 2 teachers and these evaluations were collated to obtain a single measure. In this study, an approach based on personality typologies is adopted, separating students into 3 groups according to their scores on the 5 personality traits. [26].

The study in children is more complex than the study in adults, as [4] have pointed out in their 2016 article. They stress the fact that descriptions of children's personalities by adult observers can introduces biases. These biases come from adults´ own cognitive categories and that they not always have the opportunity to observe children in different situations. On the other hand, using questions based on verbal formulations as used in youth and adults may not work well with preadolescent

children, who sometimes have not experienced the situations described. Moreover, many children of these ages have not yet fully developed their reading comprehension skills. For this reason, [4] proposes a questionnaire based on pictorial scales, according to the level of cognitive development of children. They also considerably reduce the number of items in each personality trait (just 3 questions per trait), and propose situations in each question that are specific familiar situations for children.

It is then interesting to consider [4] suggestions, using their shorter version, with questions with situations more typical for fourth grade students. In addition, their proposal could be more efficient and cost-effective if teachers administer the questionnaire to their classes and do it in less than 30 minutes as part of a regular classroom session. Therefore, we are interested in determining whether this simplified version of the Big 5 questionnaire and administered in regular classroom sessions, is adequate and manages to capture the 5 personality traits. Moreover, we are interested to know whether the big 5 model correlates with students´ math performance. It is particularly important the performance in fourth grade since several countries measure the performance at that grade level. For this reason, we decided to replicate the Polish study in Chile, but in a large-scale school setting with fourth-graders.

3. Materials and Methods

3.1. Instrument

We used the PPTQ-C (Pictorial Personality Traits Questionnaire for Children) questionnaire [4]. This is a questionnaire already validated in a study of children in Poland in 2014. It has the advantage that it has been adapted to the cognitive level of children, using specific stimuli and pictorial scales with 5 response alternatives. This is very relevant, because the options in the questions use images that represent familiar behaviors for children. Abstract linguistic statements make their understanding more difficult for children. In addition, the characters featured in each image are unisex. Therefore, the PPTQ-C is in principle suitable for both boys and girls. We translated the questionnaire into Spanish. We designed an online version using the ConectaIdeas platform [27, 28]. This educational platform uses gamification for teaching mathematics. All these classes were using the platform once a week during the semester. Teachers administered the questionnaire at the end of the school year as part of a regular math session with the platform.

During a regular math session in school hour, teachers sent each question from their smartphone or tablet to all their respective students, while monitoring whether they answered it. The teachers also projected the screen with the images, the question and the answer options on a white blackboard or projection screen. During the projection of the question, teachers read the question aloud to ensure that all students understood it correctly. They would then ask if someone did not understand correctly, and in that case teachers repeated the reading aloud. Only when teachers verified that everyone responded they did go on to the next question. The questionnaire includes an initial test question in order to familiarize students with the type of question and scale. Therefore, the total questionnaire comprised 16 items. In this initial question, students learned about the 5 options, their meaning, and how to answer by clicking. Although the students were familiar with the ConectaIdeas platform, they were not fully familiar with this 5 particular options. After the first question, the platform displayed the 15 traits-related questions. Each child responded on a personal computer or tablet, on the ConectaIdeas online platform, following the teacher's instructions. All the students were advancing in a synchronized way in each question. The average time it took each class to complete the entire activity was 20 minutes.

The great advantage of this method is that a single teacher applies the test simultaneously to his entire course (the average number of students per course is 25). The internet connection had problems in some cases. In those cases (11 schools out of 137), teachers administered a paper version. Afterwards, the teachers entered the responses into the same database. Obviously, the interesting thing about this on-line method is that it is very efficient and cost-effective, allowing the use of massive samples based on a wide range of schools. This gives it an enormous statistical power.

 doi:10.20944/preprints202308.1302.v1

However, in principle, we cannot be sure to obtain the same quality of data collection gathered in smaller studies applied to children one by one by a psychologist or dedicated teachers.

Teachers previously received a 20-minute online instruction on how to administer the questionnaire. For this, they used a google meet meeting. First, teachers answered the questionnaire as if they were students. Then they practice to command the questionnaire using ConectaIdeas platform, carry out the projections and readings out loud the questions.

3.2. Sample

We conducted the study with a sample of 3,423 4th grade students, belonging to 137 public or private subsidized schools (charter schools) in the Metropolitan region of Santiago. We did not include paid private schools in the sample. Public or subsidized schools serve 91% of the country's students. In our sample, 52% of the students attend a subsidized private school and 48% a public school (administered by the municipality or a local state education service). This is a very similar distribution to the total distribution of the country.

The selection of the schools was not 100% random, since participation in the study was voluntary. A little number of contacted schools decided not to participate. The main reason not to participate was the quality of the internet connection for the whole semester use of the platform for math sessions. Despite this, the sample is quite representative of the socioeconomic reality of 91% of the country's students, excluding the 9% who study in paid private schools and who, for the most part, belong to the highest income group in the country.

This is important, because many studies carried out on students are university students, from very specific universities and sometimes from very specific careers, who belong to the country's socioeconomic elite, or at least to upper-middle income sectors. This is a massive study carried out on students of medium-high, medium-low and vulnerable condition, with a percentage of vulnerability similar to that of the population of the Metropolitan Region of Chile (RM). Table 1 shows the vulnerability index of the sample and the one of RM. This official vulnerability data is only available for public and private subsidized schools.

Table 1. Vulnerability Index (IVE) in the sample and schools in the Metropolitan Region (RM).

Vulnerability index in sections	Sample	RM
Less than 35	--	1.1%
From 35 to 44	3.3%	1.8%
From 45 to 54	2.3%	4.4%
From 55 to 64	3.6%	5.9%
From 65 to 74	16.5%	13.2%
From 75 to 84	28.8%	25.0%
From 85 to 94	42.6%	42.5%
From 94 to 100	2.9%	7.1%
Average	81%	80%

We carried out the data collection between November 18 and December 19, 2022. In each school, we randomly selected one of the 4th grade courses to apply the survey. The total number of courses tested was 137 courses.

3.3. Non-response rate

We applied the questionnaire to 3,423 fourth-graders, but not all of them responded to the 15 items. This was not due to a lack of interest or a voluntary decision not to answer (since the ConectaIdeas platform did not allow progress to the next item if the current question was not answered), but rather to computer connection problems and instability of the Internet network in the schools. Therefore, it is reasonable to expect that the "non-response" have a random behavior.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

Indeed, 18% of the students (615 of the total) presented no response in at least one of the 15 items. Analyzing these students, we found that the vast majority had no response in only 1 or 2 items. Table 2 shows the distribution of non-response according to the number of unanswered items.

Table 2. Distribution of non-response according to the number of unanswered items.

Total number of students	3.423
Answered all the items	2.808
Did not answer at least one item	615
No response in 1 item	278
No response in 2 items	112
No response in 3 items	38
No response in 4 items	39
No response in 5 items	29
No response in 6 items	20
No response in 7 items	21

It is also important to note that the non-response rate is relatively similar for the 15 items in the questionnaire, as seen in Table 3. This relatively uniform distribution reflects the rather random nature of non-response.

Table 3. Quantity of non-response per item.

Item	**Trait**	**Direct or reverse**	**% of no response**
A1	Agreeableness	Direct	3.7%
A2	Agreeableness	Reverse	4.1%
A3	Agreeableness	Direct	4.7%
C1	Consciousness	Reverse	2.9%
C2	Consciousness	Direct	4.0%
C3	Consciousness	Reverse	4.6%
E1	Extraversion	Direct	4.1%
E2	Extraversion	Reverse	3.6%
E3	Extraversion	Direct	4.1%
N1	Neuroticism	Reverse	3.6%
N2	Neuroticism	Direct	3.4%
N3	Neuroticism	Reverse	4.0%
O1	Openness	Direct	3.6%
O2	Openness	Reverse	3.4%
O3	Openness	Direct	5.1%

Non-response distributes in a similar way by gender, as can be seen in Table 4, where we have grouped in a single section when the number of "non-response" is 5 or more. The Chi-2 statistic (calculated on the ungrouped distribution) is not significant (p=0.055) and the Phi, V-Cramer, and contingency coefficient statistics (all equal to 0.06) are very small; the nominal coefficient Eta, gives a value close to zero (0.03) for the effect of gender on the distribution of "No response". Therefore, we can conclude that there is no differentiating effect of the gender of the students for no response.

Table 4. Distribution of non-response according to gender.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

	Boys	Girls
Students answered all the items	82.1%	84.4%
Could not answer 1 item	8.9%	7.5%
Could not answer 2 items	4.2%	2.4%
Could not answer 3 items	1.2%	1.1%
Could not answer 4 items	1.1%	1.2%
Could not answer 5 or more items	2.5%	3.4%

If we look at the effect of math performance on the "No Response" distribution, we find no effect. In this case, the Chi-2 statistic is clearly not significant (p=0.89).

3.3. Treatment of non-response

Since each personality trait is constructed as an average of the 3 items used to measure it, a personality trait can be calculated for a student as long as the student has answered at least 1 item in that trait. If he does not answer any of the 3 items, that trait cannot be calculated and the student is excluded from the sample. This rule is the most lax. It leaves only 49 students out of the sample, leaving 3,374 students.

A slightly stricter rule considers only students who answered at least 2 items for each personality trait. In this case, the mean of the trait is more robust. If we apply this rule, we exclude 163 students, leaving 3,260 students.

Finally, the strictest rule can be applied, which requires that the 3 items have been answered in each personality trait, which would exclude 613 students. The remaining 2,810 have answered the 15 items of the questionnaire.

Table 5 below summarizes the 3 scenarios described:

Table 5. The 3 non-response scenarios analyzed.

SCENERY	Explanation	Number of students
They answer the 15 items	Students who answered the 3 items in each trait are left; that is, they answered the 15 items of the questionnaire.	2,810
Answer at least 2 items per trait	Students who answered at least 2 items in each trait are left. Blanks are replaced by the average of the other trait items.	3,260
Answer at least 1 item per trait	Students who answered at least 1 item in each trait are left. Blanks are replaced by the average of the other trait items.	3,374

In the first two scenarios, an additional problem is generated, since by averaging the responses of 2 items within the same range, the result may give a response not contemplated in the scale (for example, 4.5). Fortunately, there are very few cases, and the solution implemented was the following:

1.5 and 2.5 are rounded to the lower integer (1 and 2 respectively)

3.5 and 4.5 are rounded up to the next integer (4 and 5 respectively)

As an example, we present the distribution of answers for question A1 (Table 6), where there are 21 cases whose average results with one decimal place (0.62% of the total):

Table 6. Number of answers for the values of item A1.

 doi:10.20944/preprints202308.1302.v1

Value	1	1.5	2	2.5	3	3.5	4	4.5	5
Answers	303	1	266	1	378	7	760	12	1646

On average, the number of values with a decimal place is 22.4 per item (0.66% of the total per item), so the example shown for A1 is quite representative.

3.4. "Big Five" model in the different scenarios of "No response"

We applied the ESEM model using Rstudio v 4.2.3 defining the data as ordinal, using polyserial correlations and the estimation method was weighted least squares in its robust variant (WLS-MV).

The factorial model used was the same applied in the Polish study, to make the results more comparable. The main factor loadings are freely estimated, while the cross factor loadings are constrained to be small.

As a measure of model adjustment, the CFI indices (expected to be greater than 0.95) and the RMSEA and SRMR indices (expected to be less than 0.05) were considered. The Tucker-Lewis Index (TLI) is also reported [29].

We tested the model with the 3 described scenarios and the results are in Table 7.

Table 7. Adjustment of the model in the 3 "No Response" scenarios considered.

Fit test		Answer all 15 items	Answer 2 items per trait	Answer at least 1 item per trait
Respondents		2,810	3,260	3,374
CFI	Must be > 0.950	0.957	0.959	0.961
TLI	Must be > 0.950	0.933	0.936	0.939
RMSEA	Must be < 0.050	0.041	0.041	0.041
SRMR	Must be < 0.050	0.041	0.040	0.040

As can be seen, the fit of the model to the observed data is good in the 3 non-response scenarios and there are no relevant differences in the quality of the fit depending on the number of students filtered out due to non-response.

It could be thought that the more lax criterion (they answer at least 1 item per trait) favors the results, by estimating some data with the averages of the other items, thus improving the internal correlation of the items of the same trait. If so, as we see the effect is minimal. For this reason, we will take the intermediate case as the analysis scenario (at least 2 items per trait respond), to try to keep the sample as large as possible and the difference with the most demanding criteria as minimal as possible.

3.5. Chosen scenario: student demographic profile

The scenario chosen then considers 3,260 students, whose distribution by gender is as follows (Table 8):

Table 8. Distribution of the sample by gender.

Gender	Number of students	%
Male	1,700	52.1%
Female	1,560	47.9%

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

For the distribution by age (Table 9), we take the age attained on the date of application of the test. The study took place towards the end of the school year (November and December), for which the age is greater by almost one year than the age of entry of the students to 4th grade.

Table 9. Distribution of the sample by age.

Age	Number of students	% of the total	% of those who reported age
9 years	716	22.0%	23.1%
10 years	2,096	64.3%	67.7%
11 years	234	7.2%	7.6%
12 years	42	1.3%	1.4%
13 years	6	0.2%	0.2%
not reported	166	5.1%	

Finally, we show the distribution of students according to their performance in mathematics (Table 10), taking the grade point average (GPA) in math during the 2022 school year as a performance index. In Chile, the grades goes from 1 to 7, with 7 being The best grade.

Table 10. Distribution according to performance in mathematics (GPA).

GPA	Number of students	% of the total
Between 1 and 2	1	0.03%
Greater than 2.0 up to 3.0	8	0.25%
Greater than 3.0 up to 4.0	123	3.81%
Greater than 4.0 up to 5.0	676	20.92%
Greater than 5.0 up to 6.0	1315	40.70%
Greater than 6.0 up to 7.0	1108	34.29%
Not reported	29	

4. Results

We now report the results obtained in more detail, applying the factorial model already detailed in the group of students who answered at least 2 items per personality trait. First, the quality of fit obtained in our study is very similar to the study carried out in Poland (Table 11), in the RMSEA and SRMR indices. In the CFI and TLI indices, we have a slightly lower result, which we could explain by the more massive way of applying the test, as we mentioned before. It could also be the effect of the children having a slightly younger age range than the study done in Poland (children 10-13 years).

Table 11. Fit index for Chile and Poland.

Fit test		Study in Chile	Study in Poland
Respondents		3,260	527
CFI	Must be > 0.950	0.959	0.990
TLI	Must be > 0.950	0.936	0.973
RMSEA	Must be < 0.050	0.041	0.040
SRMR	Must be < 0.050	0.040	0.044

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

4.1. Individual results by item

Here we report the distribution (in %) of the responses for each of the 15 items of the questionnaire (Table 12).

Table 12. Distribution of frequencies by item (in %).

Item	1	2	3	4	5
A1	9.0	8.0	11.2	22.8	49.0
A2	2.9	2.0	8.0	15.7	71.4
A3	17.1	8.6	13.9	22.8	37.7
C1	6.4	7.5	8.1	26.0	51.9
C2	12.9	13.9	9.9	16.1	47.2
C3	14.8	5.5	7.9	11.6	60.3
E1	20.2	9.7	8.6	11.3	50.1
E2	16.5	9.7	15.0	21.1	37.7
E3	14.8	9.2	12.5	19.3	44.2
N1	63.6	10.6	11.2	7.5	7.1
N2	27.0	13.1	10.7	18.0	31.3
N3	26.7	8.2	14.4	19.6	31.1
O1	25.3	10.7	20.5	18.6	24.8
O2	36.5	4.4	8.4	5.9	44.8
O3	11.7	8.3	10.7	21.6	47.6

We also report the mean and standard deviation of each item (Table 13):

Table 13. Mean and SD of each measured item.

Item	Min	Max	Mean	SD	Var coef.
A1	1	5	3.9	1,3	0.33
A2	1	5	4.5	0.9	0.21
A3	1	5	3.6	1.5	0.42
C1	1	5	4.1	1,2	0.30
C2	1	5	3.7	1.5	0.40
C3	1	5	4.0	1.5	0.38
E1	1	5	3.6	1.6	0.45
E2	1	5	3.5	1.5	0.42
E3	1	5	3.7	1.5	0.40
N1	1	5	1.8	1,3	0.70
N2	1	5	3.1	1.6	0.52
N3	1	5	3.2	1.6	0.50
O1	1	5	3.1	1.5	0.49
O2	1	5	3.2	1.8	0.57
O3	1	5	3.9	1.4	0.36

4.2. Items and traits: reliability measure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

As we know, the proposed model assumes that each personality trait is a "latent" variable and each associated item is an observable index of that trait; for them it is necessary to analyze the reliability of each personality trait. The correlation between items within each trait is expected to be significant. The Cronbach's alpha coefficient, an index based on the average of the correlations between all the pairs of items that make up the composite variable (Table 14), measure this significance. There is no statistical test for this indicator, which varies between 0 and 1. However, the closer it is to 1, the more the measurement of the latent variable is considered to be reliable [30].

Table 14. Cronbach's alpha by personality trait.

Trait	item1 vs 2	item1 vs 3	item2 vs 3	Average correlation	Cronbach's Alpha	Alpha Polish study
Agreableness	0.38	0.37	0.34	0.36	0.63	0.67
Conscientiousness	0.42	0.29	0.28	0.33	0.59	0.61
Extraversion	0.26	0.19	0.23	0.22	0.46	0.50
Neuroticism	0.18	0.30	0.20	0.22	0.46	0.62
Openness	0.08	0.19	0.11	0.13	0.31	0.44

We obtained similar reliability results to the Polish study in agreeableness, conscientiousness, and extraversion traits. In neuroticism, we have a more noticeable difference, with a significantly lower degree of reliability than that found in Poland. In the trait of openness to experience, in both countries the degree of reliability is low. This suggests the convenience of exploring ways to improve the measurement instrument for this trait in the future.

4.3. Factor loadings

Here we show the factor loadings obtained in our study (Table 15). As expected, the cross charges are quite low, while the main charges have much higher positive values.

Table 15. Factorial loads of each item in the personality traits.

Item	**Agreeableness**	**conscientiousness**	**Extraversion**	**Neuroticity**	**Openness**
A1	0.60	0.01	0.03	-0.06	0.02
A2	0.62	0.01	0.03	-0.06	0.03
A3	0.59	0.01	0.03	-0.13	0.02
C1	0.02	0.66	0.03	-0.13	0.03
C2	0.02	0.64	0.03	-0.16	0.03
C3	0.01	0.44	0.02	-0.08	0.02
E1	0.01	0.01	0.34	-0.13	0.02
E2	0.01	0.01	0.50	-0.11	0.02
E3	0.01	0.01	0.55	0.05	0.02
N1	0.01	0.01	0.03	0.60	0.02
N2	0.01	0.01	0.02	0.47	0.02
N3	0.01	0.01	0.02	0.34	0.02
O1	0.01	0.01	0.02	0.05	0.35
O2	0.01	0.01	0.02	0.08	0.14
O3	1	5	3.9	1.4	0.36

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

Let us compare the main factor loadings with the study done in Poland (Table 16):

Table 16. Main factor loadings in the studies of Chile and Poland.

Item	Chile	Poland
A1	0.60	0.64
A2	0.62	0.56
A3	0.59	0.54
C1	0.66	0.64
C2	0.64	0.74
C3	0.44	0.58
E1	0.34	0.63
E2	0.50	0.64
E3	0.55	0.20
N1	0.60	0.64
N2	0.47	0.44
N3	0.34	0.87
O1	0.35	0.41
O2	0.14	0.14

4.4. Model with null cross factor loadings

We also test an ESEM model with all cross-loads equal to 0. This is equivalent to a confirmatory analysis model (in Rstudio the CFA function). This is a simpler model, but less flexible, so the quality of the fit could be lower, but since the cross-loads are quite close to 0, it is reasonable to expect a good fit. In general, we consider two models equivalent if the delta of the difference in CFI < 0.020 and the difference in RMSEA < 0.030. As can be seen in Table 17, both models fit the observed data well, which is why we consider them equivalent. Since when cross loading=0 the model is simpler, we will stay with this model to calculate the personality traits (latent variables).

Table 17. ESEM model and CFA model fit indices.

Fit test	ESEM model (cross loads < 0.2)	CFA model (cross loads =0)	Delta (absolute)
Respondents	3,260	3,260	
CFI	0.959	0.953	0.006
RMSEA	0.041	0.040	0.001

Table 18 shows the factor loadings. They are practically the same.

Table 18. Factor loadings for each model (ESEM and CFA).

Item	ESEM Model	CFA Model
A1	0.60	0.59
A2	0.62	0.62
A3	0.59	0.60
C1	0.66	0.65
C2	0.64	0.64

C3	0.44	0.44
E1	0.34	0.37
E2	0.50	0.52
E3	0.55	0.51
N1	0.60	0.60
N2	0.47	0.48
N3	0.34	0.33
O1	0.35	0.34
O2	0.14	0.13

4.5. Gender invariance

We tested this last model separately between men and women /Table 19).

Table 19. Adjustment of the CFA model by gender.

Fit test	Total	Boys	Girls	Delta (absolute)
Respondents	3,260	1,700	1,560	
CFI	0.953	0.951	0.965	0.014
RMSEA	0.040	0.039	0.038	0.001

The model fits well for both genders and is thus invariant by sex, since the difference in the CFI is less than 0.020 and the difference in the RMSEA is less than 0.030, which are the thresholds for accepting or rejecting invariance.

The same result was obtained in Poland (invariance of the model by gender), which confirms that the questionnaire, whose images were designed in unisex mode, worked correctly in this regard.

4.6. Invariance according to performance in mathematics

In addition, we define three groups according to their performance in mathematics, taking as an indicator the GPA:

- Very good: students with a math average above 6.0
- Good: students with an average in mathematics between 5.01 and 6.0 (both inclusive)
- Regular: students with an average in mathematics equal to 5.0 or less

Excluding students without qualification (29 cases, 0.9% of the total), the model was tested separately in the 3 groups. Table 20 shows the fit of the model in each group:

Table 20. Adjustment of the model according to performance in mathematics.

Fit test	Total	Very good	Good	Regular	Delta (absolute)
Respondents	3,260	1,108	1,315	808	
CFI	0.953	0.966	0.954	0.949	0.017
RMSEA	0.040	0.034	0.042	0.039	0.008

We can then observe that the Big Five model is invariant according to the children's performance in mathematics. This gives us peace of mind in the sense that understanding the questionnaire was not a problem that could affect the results of the study. This conclusion is based on the solid evidence

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

that exists in the sense that there is a strong correlation between performance in mathematics and reading comprehension [31].

4.7. Analysis of personality traits

Each personality trait is a latent variable and is calculated as a weighted average of the 3 items measured in the questionnaire. The normalized factor loads gives the weighting factor that the CFA model shows (in this case, the sum of the coefficients is 1.0). Table 21 shows the calculation.

Table 21. Normalized coefficients for face personality trait.

Item	Factor loadings	Normalized coefficients
A1	0.59	0.328
A2	0.62	0.341
A3	0.60	0.331
C1	0.65	0.375
C2	0.64	0.370
C3	0.44	0.255
E1	0.37	0.264
E2	0.52	0.371
E3	0.51	0.365
N1	0.60	0.425
N2	0.48	0.339
N3	0.33	0.236
O1	0.34	0.313
O2	0.13	0.120

Performing this calculation for each personality trait, we can estimate the frequency distribution for each trait (Figure 1).

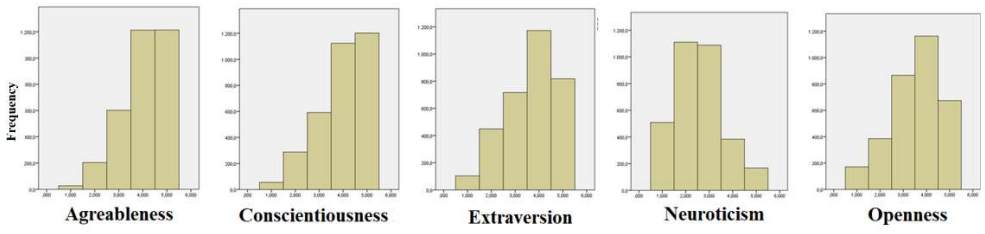

Figure 1. Distribution of frequencies for each personality trait.

Taking the n=3,260 cases, where we have calculated the values for the 5 personality traits, we have the following averages and measures of variability. In all the variables the minimum value is 1 and the maximum is 5 (Table 22).

Table 22. Mean and SD of the 5 personality traits (Chile study).

Study in Chile 4th graders 9-12 years old	**Agreeableness**	**conscientiousness**	**Extraversion**	**Neuroticity**	**Openness**
Mean	4.0	3,9	3.6	2.6	3.5
SD	0.88	0,98	1.02	0.99	1.02

 doi:10.20944/preprints202308.1302.v1

Var Coef.	22%	25%	28%	38%	29%

Table 23 shows the means by personality trait in the study from Poland (children 10-13 years old):

Table 23. Mean and SD of the 5 personality traits (Poland study).

Study in Poland 10-13 years old	**Agreeableness**	**conscientiousness**	**Extraversion**	**Neuroticity**	**Openness**
Mean	3.1	3.6	3,1	2.4	3.2
SD	0.87	1.01	0,87	1.01	1.05
Var Coef.	28%	28%	29%	42%	32%

We can then see that the averages in the Chilean study are slightly higher than those obtained in the Polish study, but in both cases, the neuroticism trait is the one that obtains the significantly lower average (compared to the other personality traits). The standard deviations are quite similar in both studies.

Lastly, we report the correlation between the five personality traits (Table 24). The calculated correlation is Pearson's, since each trait is a weighted combination of variables, taking many tens of different values.

Table 24. Correlation of latent variables (2 item x trait).

Trait	Agreableness	Conscientiousness	Extraversion	Neuroticism	Openness
Agreableness	1.00				
Conscientiousness	0.33	1.00			
Extraversion	0.25	0.16	1.00		
Neuroticism	-0.17	-0.24	-0.22	1.00	
Openness	0.35	0.31	0.22	-0.17	1.00

The correlation between traits is low or moderately low, which gives consistency to the "Big Five" model. In additon, the correlation of neuroticism with the other traits is negative.

4.7. Personality traits and gender

We calculated the average of the 5 personality traits for each gender and differences of more than 2 tenths (statistically significant) were observed in agreeableness and neuroticism, with higher values in girls than boys in both cases (Table 25).

Table 25. Personality traits and gender.

Student Gender	**Agreeableness**	**conscientiousness**	**Extraversion**	**Neuroticity**	**Openness**
Male	3.9	3.9	3.7	2.5	3.5
Female	4.1	3.9	3.6	2.7	3.6
Significance (ANOVA)	0.000	0.229	0.005	0.000	0.000

In the Polish study, girls also scored higher on the neuroticism trait, but they also scored higher on openness to experience. In our study, we also detected this difference in favor of women, but only with a difference of one tenth (statistically significant).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

In Chile, we also found a small difference (1 decimal, but statistically significant) in favor of men in the extraversion trait.

4.7. Personality traits and performance in mathematics

A subject that has been widely studied is the relationship between personality traits and cognitive abilities, particularly performance in so-called "quantitative" subjects (mathematics, science, technology, etc.). There is plenty of empirical evidence that a greater degree of openness and conscientiousness positively influence cognitive abilities [19] and better school performance [22]. In our study, we analyzed performance in mathematics, measured through the grade point average (GPA) throughout the school year. Figure 2 shows the distribution of standardized math GPAs for 4th grade students (n=3,231 children, since there are 29 students without a reported grade).

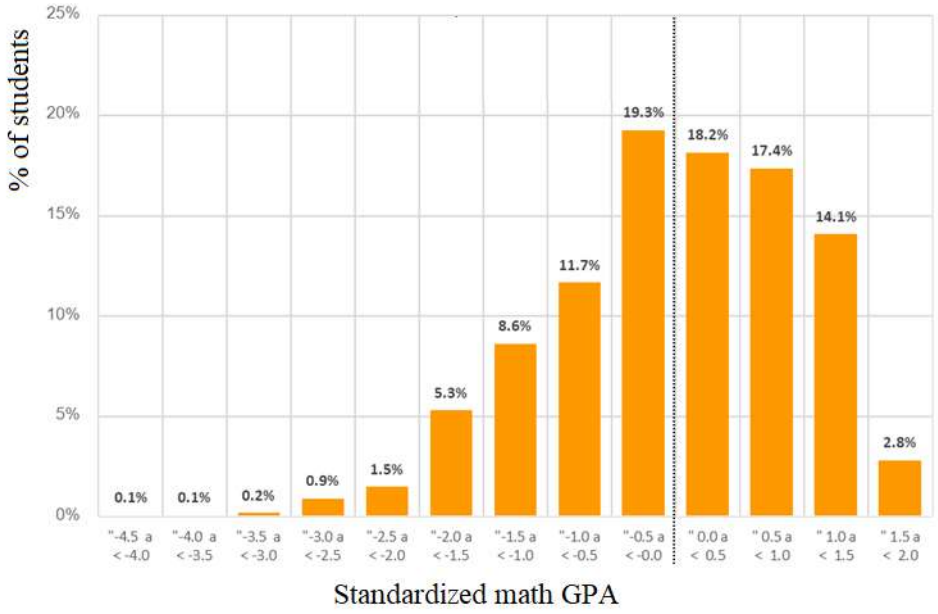

Figure 2. Distribution of math GPAs.

We calculated the correlation between math performance and personality traits. We found statistically significant (95% confidence), but weak relationships between performance in mathematics and the traits of extraversion and openness to experience, as shown in Table 26:

Table 26. Personality traits and performance in mathematics.

Pearson Correlation	**Agreeableness**	**conscientiousness**	**Extraversion**	**Neuroticity**	**Openness**
Performance	-	-	0.06	-	0.04
Significance (ANOVA)	0.303	0.977	0.000	0.082	0.000

This poor result is partly due to the great concentration of the data in a few values in the personality traits. We can improve the analysis by taking an alternative approach, focusing on people rather than variables, looking for groups of students with similar personality profiles that can account for differences in mathematics performance. Other researchers have followed this line of typological analysis, such as the aforementioned study carried out in China by [26].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

If we divide students into 3 groups for each personality trait (those who score high on the trait, those who score low, and those in between), differences in math performance are seen in 3 of the 5 traits: conscientiousness, extraversion, and opening (Table 27):

Table 27. Mathematics performance (average grade) by group according to personality trait.

Score on each trait	Agreableness	Conscientiousness	Extraversion	Neuroticism	Openness
Low	-.036	-.103	-.128	.033	-.113
Intermediate	.031	.067	.006	.006	.011
High	-.017	-.011	.067	-.088	.037
High-low difference	0.066	0.170	0.196	0.121	0.150
Anova test	0.352	0.004	0.000	-0.062	0.009

The Table shows a significant difference in mathematics performance between the groups with low scores in each of the conscientiousness, extraversion and openness traits, compared to the groups with medium and high scores. This difference is on average 0.14 points (on a scale of 1 to 7) and in terms of standard deviations (0.82 for the total sample), it is equivalent to a difference of 0.17 standard deviations.

The influence of conscientiousness and openness on cognitive abilities has been widely reported [19- 22]. Some studies have also found a positive influence of the 3 personality traits indicated (conscientiousness, openness, extroversion) on academic self-efficacy [32].

4.8. Personality types and performance in mathematics

One wonders if combining the different personality traits can further differentiate performance in mathematics. That is, if students who score high simultaneously on 2 or more traits could still perform better than those who score high on only one trait. We performed a cluster analysis to build a segmentation with the 5 traits simultaneously, but the result failed to improve what was already presented above.

We then centered our analysis on the different possible combinations between the 3 traits that do discriminate in performance in mathematics (conscientiousness, extraversion and openness) and the best result is found by segmenting a common index for these 3 variables. This common index is a weighted average of the aforementioned traits. We obtained this as a solution of the factorial analysis in principal components. This solution precisely has the mathematical property of building a new variable (or index), weighting each one of the 3 traits, in such a way that the variance of the index is the maximum possible. In this way, students can be better discriminated according to their personality traits. Table 28 presents the result of this analysis.

Table 28. Extraction of factors by Principal Component Analysis.

Trait	1	2	3
Extraversion	.61	.77	.17
Openness	.77	-.18	-.62
Conscientiousness	.72	-.47	.52
% of variance explained	49%	28%	23%

We then performed a cluster analysis (using K-Means in SPSS) and detected 5 clusters in this index, from the lowest common score to the highest common score. Table 29 shows the size and relative position of each group and Figure 3 does it in graphical form.

Table 29. Size and position of each group in the First Principal Component.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

Conglomerate	G1	G2	G3	G4	G5
1st main component	-2.11	-1.04	-0.22	0.51	1.31
Number of cases	215	632	879	881	653
% of cases	6.6%	19.4%	27.0%	27.0%	20.0%

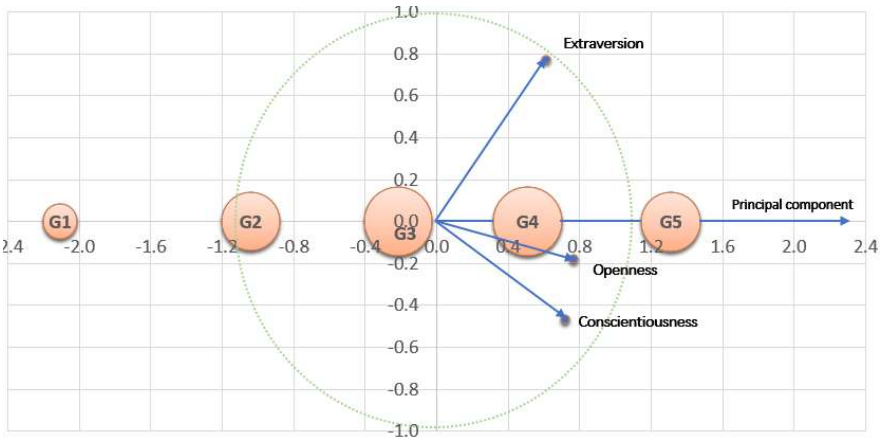

Figure 3. Representation of the variables and clusters in the plane defined by the first 2 factors.

Finally, we made a comparison of the average for mathematics in each of these 5 groups (using ANOVA in SPSS), resulting in a highly significant difference between the averages (p=0.003). Figure 4 shows the variation of the mathematics average in each group. We see that the G1 group, which has the lowest score on the common factor of conscientiousness, openness to experience, and extroversion, clearly also has the lowest average in mathematics. The last 3 groups (G3, G4 and G5) are those with the highest average in mathematics, with no significant difference between them.

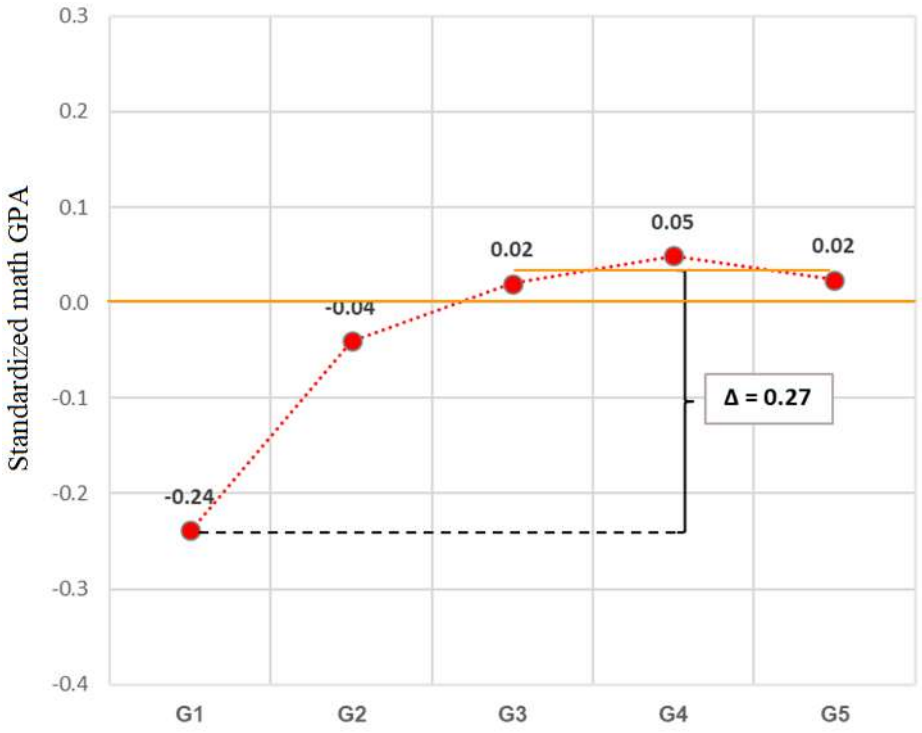

Figure 4. Mathematics performance according to the 5 groups defined by conscientiousness-openness-extraversion.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2023 doi:10.20944/preprints202308.1302.v1

The important finding here is that the difference between the group with the lowest performance in mathematics (G1: 6.6% of the students) and the average of the groups G3, G4 and G5 that have a high performance, but without statistical difference between them (74.3% of students), is 0.27 standard deviations. This is a very significant difference.

In Table 30, we report the average of each personality trait in the 5 groups already defined. We can see that the three traits systematically increase their average from group G1 to group G5. This allows Group 1 to be interpreted as the group of students with the lowest score simultaneously in the 3 personality traits and thus growing, up to Group 5, which are the students with the highest score in the same 3 traits.

Table 30. Average of each trait by personality group.

Group	N	Conscientiousness	Openness	Extraversion
G1	215	2.3	1.8	2.5
G2	632	3.2	2.7	3.0
G3	879	3.8	3.4	3.4
G4	881	4.3	4.0	3.9
G5	653	4.8	4.5	4.5
Total	3260	3.9	3.5	3.6

Finally, we highlight that performance in mathematics does not grow linearly with the increase in socio-emotional skills measured in factor 1 (first principal component of conscientiousness, extraversion and openness traits). Performance grows rapidly at first, but quickly levels off. This indicates that teachers and school psychologists should focus on supporting students who have serious deficits in these skills. This psychological support would help improve their performance in mathematics. There is considerable evidence that early intervention in children, where their personality traits are still somewhat malleable, pays off, improving not only their cognitive abilities, but also many aspects related to their satisfaction with life [20].

5. Discussion

One objective of this study was to replicate in Chile the study carried out in Poland, using the PPTQ-C (Pictorial Personality Traits Questionnaire for Children) questionnaire, to test its validity in a larger sample, with administration of the questionnaire at regular classroom sessions, and applied by the same teachers of the students.

Indeed, teachers applied the questionnaire in the classroom under normal classroom conditions and in their regular math sessions. The students are from public or private schools subsidized by the state. Thus, they are not students from high social strata or from the cultural elite of the country. On the contrary, they are students who represent the vast majority of students of medium and low socioeconomic status, including students in vulnerable and neurodivergent conditions.

Therefore, the method used in this study has the advantage of being easily scalable at low cost. A possible disadvantage comes from this same massiveness and "natural" application methodology. There was concern that some students had problems understanding the survey. There was also a concern of loss of responses due to technical problems of the computers, the Wi-Fi networks and the internet connectivity.

However, the results show that the study nicely overcame those potential drawbacks. Non-response has no impact on the results, since it is rather random than biased. On the other hand, the fact that the questionnaire was applied synchronously by the teacher, who read each question and waited for the students to answer it (on a computer or tablet, online), helped to overcome possible reading comprehension problems. As we have seen, the model fits well top, intermediate and low performers in mathematics.

We found that this study supports the validity of the Big Five model in children at this school level (4th grade, children between 9 and 11 years of age). The study allows us to recognize the 5 factors, although with different levels of reliability. Similar to the study carried out in Poland, the openness trait obtained a lower level of reliability. This suggests that it is convenient in the future to improve the measurement instrument for this trait.

We also independently tested the model on both genders and found a good fit for both boys and girls (as in Poland), which validates the unisex nature of the ideograms used. Since the questionnaire works well in both genders, we checked if there are differences in personality traits between both genders. We obtained a small but statistically significant difference. This finding is in line with the findings of other studies in this age range. Girls score slightly higher on the traits of agreeableness and openness, while boys score slightly higher on emotional stability (the opposite of neuroticism) and on extroversion. We did not detect differences in the conscientiousness trait.

We found an influence of conscientiousness, openness to experience, and extroversion traits on performance in mathematics, measured with the GPA obtained during the whole year. While each of these personality traits exerts an influence on its own, this impact is doubled if we consider all 3 traits in combination. The best way to do it is to take an index (weighted average of the 3 variables) that has maximum variance. This is exactly the first principal component of the factorial analysis applied to the 3 variables mentioned. If we classify students into 5 groups (G1, G2, G3, G4 and G5) according to this index (for example using SPSS K-Means), we obtain a clear difference in performance in mathematics. The group with the lowest scores in the conscientiousness, openness and extraversion (G1) personality traits, has the lowest performance in mathematics. The second group (G2) has a medium performance. The group formed by the union of G3, G4, G5 has the highest math performance, without significant differences between them. The difference in the math GPA between G1 and the average of G1-G2-G3 reaches 0.27 standard deviations. This is a big difference.

The previous result has several practical aspects of interest. It shows that the relationship between personality traits and performance in mathematics is not linear, nor symmetrical in the positive and negative directions. The group with low scores in the 3 personality traits has very low math GPA, quite distant from the rest. This shows that the deficit in these personality traits significantly affects certain cognitive abilities. However, as this combination of these three traits increases the corresponding math GPA hits a threshold. This is reached somewhere in the middle of the personality index. Going up from there no longer improves math performance. This is not surprising, as there is some evidence to suggest that extremely high levels on these personality traits can have counterproductive effects on academic performance [33, 34]. However, this finding shows the importance of carrying out an early intervention to improve the socio-emotional skills of these children with low scores in the indicated personality traits (conscientiousness, extroversion and openness). At that age, these traits are still malleable to some degree. This would result not only in improving their cognitive abilities, but also in many other aspects related to their personal satisfaction.

6. Conclusions

In this work, we have investigated whether it is feasible to carry out a large-scale measurement of the 5 personality traits of the Big 5 model, using a shorter, pictographic and appropriate version for fourth-graders. We have done it with the PPTQ-C questionnaire (Pictorial Personality Traits Questionnaire for Children) that has only 15 questions, where the student answers on an online platform, in a regular classroom session. Thus, we have replicated the previous study in Poland. We replicated it with 3,423 4th grade students (9 to 12 years old). The teachers, in regular sessions, applied the questionnaire to all of their respective classes. We found similar results, including that asexual pictograms worked well across both genders.

We then studied whether there is a relation between the Big 5 personality traits and math performance, and how each personality trait influences students' math GPAs. We obtained three findings. First, we found that there is a relation between conscientiousness, openness, and extraversion, with mathematical performance. Second, we found that a combination of these three

traits has a relation with math performance that is the double of the relation of each trait alone. Third, we find that students with the lowest scores on this combination of personality traits (6.6% of students) have 0.27 standard deviations less in math performance than those with the highest scores, which are 74.3% of students.

To the best of our knowledge, this is the first time that a study has found a strong relationship between a combination of personality traits collected with a 15-question questionnaire and fourth graders' math performance measured with GPAs.

For future work, it seems very important to us to continue accumulating evidence of the "Big Five" model in children, using pictographic and metric scales that are in accordance with their level of cognitive development. The results of this study provide arguments in favor of continuing along this path. It is especially interesting to carry out studies that are more massive and in normal classroom conditions, despite the limitations that this implies in terms of the quality of data collection.

A proposal for future work is to improve the reliability of the applied questionnaire, which, as we have seen, presents acceptable alpha values, but is still a little low. For example, add one item to each personality trait. This would make it possible to give more reliability to the measurement of each trait. It will improve the Cronbach's alpha in all of the traits, obviously under the assumption that the added items correlate correctly within each personality trait. The length of the questionnaire would increase from 15 items to 20 items, which would only imply a few more minutes in the application of the questionnaire. The analytical benefit could be much greater than the cost incurred in increased application time.

Author Contributions: Conceptualization, R.A.; methodology, R.A., P.G.; software, P.G.; validation, R.A., P.G.; formal analysis, P.G..; investigation, R.A., P.G.; resources, R.A..; data curation, R.A., P.G.; writing—original draft preparation, P.G.; writing—review and editing, R.A..; visualization, P.G..; supervision, R.A..; project administration, R.A..; funding acquisition, R.A.. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Chilean National Agency for Research and Development (ANID), grant number ANID/PIA/Basal Funds for Centers of Excellence FB0003.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to it being a class session during regular school time. The activity was revised and authorized by the respective teachers.

Informed Consent Statement: Informed consent was obtained from all teachers according to the provisions of 18.1 on the protection of data of minors of the Children's Online Privacy Protection Rule (COPPA), Section N.

Data Availability Statement: Not available.

Acknowledgments: WE acknowledge support from the Chilean National Agency for Research and Development (ANID), grant number ANID/PIA/Basal Funds for Centers of Excellence FB0003. In addition, we acknowledge technical support to Paulin Jaure and Carlos Aguirre.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Engler, B.; Velásquez, J.; Mancilla, B. *Introducción a las Teorías de la Personalidad*, Mc Graw Hill 4ª edición, 294-318. 1996.
2. McCrae, R.; Costa, P. Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*, **1987**, Vol. 52, No. 1, 81-90.
3. Heine, S.; Buchtel, E. Personality: The Universal and the Culturally. *Annu. Rev. Psychol.* **2009**, *Volume 60*, 369–394.
4. Maćkiewicz, M.; Cieciuch, J. Pictorial Personality Traits Questionnaire for Children (PPTQ-C)—A New Measure of Children's Personality Traits. *Frontiers in Psychology, Sec. Developmental Psychology*, **2016.** Volume 7.
5. Soto, C.; John, O.; Gosling, S.; Potter, J. Age Differences in Personality Traits From 10 to 65: Big Five Domains and Facets in a Large Cross-Sectional Sample. *Journal of Personality and Social Psychology*. **2011**, Vol. 100, No. 2, 330–348
6. Barbaranelli, C.; Caprara, G.; Rabasca, A.; Pastorelli, C. A questionnaire for measuring the big five in late childhood. *Personality and Individual Differences*, **2003**. Volume 34, Issue 4, Pages 645-664.

7. Del Barrio, M.; Carrasco, M.; Holgado, F. Análisis transversal de los cinco factores de personalidad por sexo y edad en niños españoles. *Revista Latinoamericana de Psicología* **2006**, volumen 38, No 3, 567-577
8. López-Cassà, E.; Pérez-Escoda, N.; Alegre, A. (2022). The Relationship between Children's Trait Emotional Intelligence and the Big Five, Big Two and Big One Personality Traits. *Education Sciences*, **2022**, 12, 491.
9. Baker, M. Is There a Reproducibility Crisis? **2016**. *Nature*. Vol 533
10. Aarts, A.; et al; and LIN, Stephanie C.. Estimating the reproducibility of psychological science. **2015.** *Science*. 349, (6251), 943-950
11. Ioannidis, J. Why Most Published Research Findings Are False. **2005.** *PLoS Medicine*. Volume 2, Issue 8, e124
12. Grossmann, M. *How Social Science Got Better*. Oxford University Press. N.Y. 2021.
13. Asendorpf J. B. & van Aken, M. A. Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: replicability, predictive power, and the trait-type issue. *Jornal Pers Soc Psychol* , **1999** Oct, 77(4):815-32
14. Twenge, J.; Cooper, A.; Joiner, T.; Duffy, M.; Binau, S. Age, Period, and Cohort Trends in Mood Disorder Indicators and Suicide Related Outcomes in a Nationally Representative Dataset, 2005–2017. *American Psychological Association*. **2019**. Vol. 128, No. 3, 185–199
15. Rutkowski D.; Rutkowski L. (2018) How systemic is international bullying and what relationship does it have with mathematics achievement in 4th grade? IEA, Amsterdam.
16. Watson A. *Care in mathematics education*. Palgrave Macmillan, London. 2021.
17. Chaudhary, N.; Swanepoel, A. Editorial Perspective: What can we learn from hunter gatherers about children's mental health? An evolutionary perspective. *Journal of Child Psychology and Psychiatry*. **2023.** ( ) 1-4 doi:10.1111/jcpp.13773
18. Araya, R. Unraveling a Royal Road to Math Education. *Constructivist Foundation*. **2023**. Vol 18. Num 2. 283-286 https://constructivist.info/18/2
19. OCDE (2019). The Ocde Survey of Social and Emotional Skills. "Beyond Academic Learning: first results from the survey of social and emotional skills"
20. OCDE (2018) Education Working Paper No. 173: Social and Emotional Skills for Student Success and Well Being: Conceptual Framework for the Oecd Study on Social and Emotional Skills.
21. Stanek, Kevin C. & Ones, Deniz S. Meta-analytic relations between personality and cognitive ability. *Proceedings of the National Academy of Sciences* **2023.** vol 120 N° 23, 1-12.
22. Meyer, J.; Jansen, T.; Hübner, N.; Lüdtke, O. Disentangling the Association Between the Big Five Personality Traits and Student Achievement: Meta-Analytic Evidence on the Role of Domain Specifcity and Achievement Measures. *Educational Psychology Review.* **2023**. 35:12.
23. Roberts, B. W.; Walton, K. E.; Viechtbauer, W. Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, **2006.** 132(1), 1–25.
24. Resing, W.; Bleichrodt, N.; Dekker, P. Measuring Personality Traits in the Classroom. *European Journal of Personality*. **1999**. 13: 493-509 (1999)
25. Zhang, J. R. (2011). A Follow-Up Study of the Structural Evaluation of Children's Personality and Its Developmental Characteristics. China: Liaoning Normal University.
26. Yu, Y.; Zhang, Y. Personality and Developmental Characteristics of Primary School Students' Personality Types. *Frontiers in Psychology*, **2021**. vol 12, article 693329.
27. Araya, R. & Van der Molen, J. (2013) Impact of a blended ICT adoption model on Chilean vulnerable schools correlates with amount of on online practice. Proceedings of the Workshops at the 16th International Conference on Artificial Intelligence in Education AIED 2013. Memphis, USA, July 9-13, 2013. http://ceur-ws.org/Vol-1009/0603.pdf
28. Araya, R.; Diaz, K. Implementing Government Elementary Math Exercises Online: Positive Effects Found in RCT under Social Turmoil in Chile. *Educ. Sci*. **2020**, 10, 244. https://doi.org/10.3390/educsci10090244
29. Hu, L.; Bentler, P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, Structural Equation Modeling: *A Multidisciplinary Journal*, **1999.** Volume 6, - Issue 1
30. Nunnally, J.C & Berstein, I.J. (1995) *Teoría Psicométrica*. 3ª Edicion Mc Graw Hill, 259-262.
31. Ünal , Zehra E & Greene, Nathaniel R. & Lin , Xin & Geary, David C. What is the Source of the Correlation Between Reading and Mathematics Achievement? Two Meta-Analytic Studies. *Educational Psychology Review*. **2023**
32. Domínguez-Lara, S.; Navarro-Loli, J.; Prada-Chapoñan, R. Ítem único de autoeficacia académica: evidencias adicionales de validez con el modelo Big Five en estudiantes universitarios. *Avaliação Psicológica*, **2019**, 18(2), pp. 210-217.
33. Jensen, M. Personality Traits, Learning and Academic Achievements. *Journal of Education and Learning*; **2015**. Vol. 4, No. 4; 2015
34. Martínez De Ibarreta, C.; Redondo Palomo, R.; Rua Vieites, A.; Fabra Florit, E. Factores de personalidad (Big Five) y rendimiento académico en asignaturas cuantitativas de ADE. *Anales de Asepuma*, XIX Jornadas Valencia 2011.

 doi:10.20944/preprints202308.1302.v1