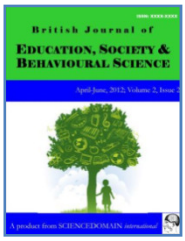

British Journal of Education, Society & Behavioural Science
4(9): 1234-1248, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Formal-Logic Development Program: Effects on Fluid Intelligence and on Inductive Reasoning Stages

Cristiano Mauro Assis Gomes[1*], Hudson Fernandes Golino[2,3], Mariana Teles Santos[4] and Michele Gomes Ferreira[5]

[1]Department of Psychology, Neuroscience Post-Graduate Program, Universidade Federal de Minas Gerais, Minas Gerais, Brazil.
[2]Neuroscience Post-Graduate Program, Universidade Federal de Minas Gerais, Minas Gerais, Brazil.
[3]Faculty Northeast Independent, Brazil.
[4]Department of Psychology, Universidade Federal da Bahia, Bahia, Brazil.
[5]Psychology Post-Graduate Program, Universidade Federal de Minas Gerais, Minas Gerais, Brazil.

Authors' contributions

This work was carried out in collaboration between all authors. All authors read and approved the final manuscript.

Original Research Article

Received 8th April 2014
Accepted 12th May 2014
Published 27th May 2014

ABSTRACT

The Formal-Logic Development Program (FLDP) is an intervention that focuses in the argumentative abilities and aims to change and increase the intelligence level of people. This study showed the first empirical analysis of the effect of the program on fluid intelligence and on the inductive reasoning stages. A 15 years old Brazilian teenager, enrolled in an education program for youth and adults, participated in the intervention (case) and the control group (N=12) was composed by teenagers enrolled in the same educational program. Four fluid intelligence tests were administered at pre and posttest. The intervention lasted approximately 4 months, 2 hours per session. The fluid intelligence score was generated using the Rasch model and the participants answers to the items of the four tests used. Regarding the adjustment of the items to the model, on average the

**Corresponding author: Email: cristianomaurogomes@gmail.com, cristianogomes@ufmg.br;*

items presented an InFit of .98 and a standard deviation of .34. Analyzing the performance of the case and the control group, both showed improvement in the posttest performance compared to the pretest. The performance of the case in the pretest was .22 logits, while the control group had an average performance of -.85 (SD =1.04). In the posttest the case achieved a performance of 3.16 logits, and the control group had an average performance of .18 (SD=1.03). The case had a gain of 2.94 logits and the control group had an average gain of 1.03 logits (SD=1.04). The effect size was 1.84 (.88 to 2.76, t=1.76, = .05 - one-tailed), when comparing the gain of the case in relation to the gain obtained by the control group. The case showed a significant increase on fluid intelligence and was the only one who presented the increment of one stage on inductive reasoning. These initial evidences indicated the effect of FLDP in the analyzed variables and that the program apparently is capable to increase intelligence. Generalizing, the results point to the direction that the flow of intelligence development can be intentionally changed by psycho-educational interventions.

Keywords: Cognitive intervention; intelligence; adult education; teenager education, formal-logic development program.

1. INTRODUCTION

It is important for any nation concerned with its people to think about how to develop their cognitive abilities and their ability to learn. In Brazil, there is a need for conceptual and methodological discussions about how to better teach students, in order to change the flow of their intellectual development and ability to think and learn. Currently, what is really new in Brazilian teaching proposals is the need for a differentiated perspective on the educational process, especially with regard to the learning concept adopted. The mainstream Brazilian teaching is still guided by the content-paradigm, which believes that education is composed by a succession of cumulative and ordered information to be transmitted to students [1]. A change in this paradigm requires the teacher to be responsible for new and insightful ways to drive the pedagogical process, which must develop mental abilities directed to the autonomy of thought, judgment and creativity, contributing for the development of conscious and critical man in the society we live in [2]. In Brazil there are great challenges to promote the complete schooling of the population and they reveal problems in the structure and methodology of the educational system. In 2011, only 26% of the population aged between 15 and 64 years old could be considered fully literate and this percentage remains the same since 2000 [3]. This educational indicator [3] reinforces the importance that education has to improve the level of literacy in the population and contribute to the debate about the quality of the educational system. Increasing access to education is necessary, but it is also necessary that the school fulfills its role to improve students' cognitive abilities. The increased access to education is not followed, most of the time, by a significant improvement in the academic proficiency. Despite the negative conditions of the Brazilian education, there are possibilities for improvement of the intelligence levels of the individuals, although little is still known about the possibility of changing it. According to Jensen [4], this has been the topic of increased investment in research in recent years in Psychology, especially in the United States, because of the importance that intelligence has to predict academic achievement, educational attainment, income level, occupational status, general health, longevity, job performance, emotional stability and crime, among other phenomena [5]. Until recently, part of the academic community concluded that the existing interventions had little or no success to modify intelligence [6]. However, during the last years, a significant number

of studies have been published showing that some interventions have a positive impact on intelligence [7].

Some cognitive programs have shown positive results in relation to the capacity of students to learn and think. The Feuerstein Instrumental Enrichment (FIE) Program is an example of such programs. Romney and Samuels [8] did a meta-analysis about the effects of FIE from 1979 to 2001, including 47 studies, and found evidences that the program has an effect on different cognitive abilities and on school domains

According to Cohen [9], the effect size is considered small when it is between .20 and .49, moderate between .50 and .79 and large when values are over .80. Both the un-weighted and the weighted effect size indicate the strength of the changes, but differently from the former, the latter is sensible to the number of studies used in its computation. Analyzing the weighted *d* obtained from Romney and Samuels [8] no construct presented a large change after the intervention, and only the verbal ability showed a moderate effect size ($d_{weighted}$=.56, p<.01). This scenario changes when we consider the un-weighted *d.* Planning-organization presented a large effect size (d=1.04, p<.01) while verbal ability (d=.61, p<.01), spatial-perceptual ability (d=.58, p<.01) and reasoning ability (d=.61, p<.01) showed moderate effect sizes. A small effect size was observed for the overall academic performance (d=.41, p<.01), mathematics (d=.35, p<.01), reading (d=.43, p<.01), artistic language (d=.47, p<.01), general cognitive ability (d=.49, p<.01), non-verbal ability (d=.41, p<.01), numeracy (d =.28, p<.05), metacognition (d=.49, p<.01), adaptative functioning (d=.28, p<.01) and locus of control (d=.20, p<.05).

Meanwhile, some studies published after 2001 have also shown evidence that the FIE has positive effects in fluid intelligence. Kozulin et al. [10] showed effect sizes ranging from .11 to .46 (p<.05) for some WISC's fluid intelligence subtests, and an effect size of .37 (p<.05) for the Raven test in children with mental age between 5-7 years old. In other study, Kozulin [11] showed a non-significant moderate effect size on the Raven's score after 120 hours of intervention (d=.45, p=.18). Studies like the one from Lizarraga, Ugarte, Cardelle-Elawar, Iriarte and Baquedano [12] that use FIE and other interventions method will not be presented here.

However it is important to highlight that the results of interventions using the FIE are not consensual. Shiell [13] in his meta-analysis, examined the efficacy of this program to increase the reasoning capacity of individual and included 36 studies from 1979 to 1996. The effect sizes of the interventions showed very inconsistent results: for verbal abilities the size effect was significant and varied from .24 to 1.41; for visual perception and visual-motor capacity, the size effects were between .42 and 1.68; for mathematical ability the effect size was .26 to .29. The author did not find a significant size effect for reading or emotional domain, like self-confidence and self-concept. The size effect for motivation was negative, which indicates that the control group outperformed the experimental group in this construct.
Due to the controversial and inauspicious results, there is a conflict in the scientific community regarding the possibility to modify the intelligence, despite the intense financial effort of international agencies and the private sector. On one side there is the development of theories like Feuerstein' structural cognitive modifiability, and as a consequence, the development of programs such as FIE. On the other hand, some researchers are very skeptic about the possibility to modify the intelligence levels. These theorists argue that because the intelligence is one of the most stable human characteristics across the life cycle, it would be impossible to change it. Moreover, they argue that the classical training programs succeed when they target specific cognitive abilities, but have no effect on general

intelligence. Another major criticism of the cognitive intervention programs and their scope refers to the difficulty of generalization and temporal stability of the results of the intervention, which usually disappear almost completely with the time [14].

It is clear, however, that new strategies to promote an acceleration of intellectual development have been developed, using well-defined interventions, focused on mediated learning. The FIE program, for example, intervenes directly in cognitive abilities and has characteristics such as: a well-defined intervention methodology, large intervention time (at least one year) and focus on metacognitive processes of analysis of the persons own thoughts and learning process [2]. Despite the challenges to its implementation, such programs, like FIE, show that it is possible to change the intelligence and the students' ability to learn. Another initiative that aims to change the intelligence development flow of people is the Formal-Logic Development Program (FLDP), developed in Brazil by Gomes [15]. The FLDP follows key principles present in the FIE program, but the former is different because it intervenes specifically on cognitive abilities at the abstract and formal developmental stages. In this sense, the FLDP works where the FIE program stops, being complementary programs. Based on this overview, the present study was motivated by the interest in investigating the effects of the Formal-Logic Development Program (FLDP) on the mental performance of young and adults, providing important results to boost the field of studies of cognitive intervention. The development of abstract and formal stages on certain cognitive abilities is an actual demand from the post-industrial society. The FLDP aims to improve abstract and formal stages, and the present study investigates if its efficiency through a single case study with a control group method.

2. INVESTIGATION ISSUE

As described above, there is an urgent need for the development of the Brazilian population interpretation, logic and argumentation abilities. Mainly for people aged between 15 and 64 years. This is a critical national problem. As it was also shown previously, some cognitive programs can improve cognitive ability and school performance, for example the FIE.

In 2003, Gomes [15] developed in Brazil the Formal-Logic Development Program (FLDP), in portuguese ´Programa de Desenvolvimento da Lógica-Formal´(PROLOF).The program uses the theoretical and methodological basis from FIE, but focuses exclusively on the intervention of the formal-abstract abilities. The program has an innovation in relation to the FIE and other programs, as it has a scale in which the performance of the student can be monitored by the instructor and the student himself. There is a scale that allows the instructor to score all the activities of the student in relation to his capacity to represent the information available in the activity and to represent the thinking process. The student also learns to use this scale and can check his progress across time.

There are five units in the program. Units I and II are composed by inductive and deductive logical reasoning. During this two units, the instructor works to develop the logical thinking ability of the student, through the representation of the information and the representation of the thinking process. During unit III the student will identify the argumentative structure in various texts, adding the logical thinking to the textual analysis. The unit IV complements unit III and now the student has to identify contradictions in arguments. Unit V is the last one and the student elaborates his own argumentative text, instead of identifying argumentative structures in already existing texts. During all the units the intervention happens through the development of the student´s ability to represent the information and his process of thinking, and also his process of monitoring his own progress using the scale.

Since the first session the basic structure of the program is presented and it is focused on the representation of the information and thinking. The instructor shows the scale and teaches the student how to fill in the scale based on how he performs in each activity, therefore the student learns how to fill it in slowly and under the instructor´s supervision. The starting point of every activity involves doing the activity and also representing graphically the information and the thinking. Scoring a task includes not only doing what is required, but also to get the greatest grade in the scale in both the information and thought representation. The FLDP can be administered individually, or in groups of up to six students, from school year 7 and on. The number of participants must be small, because it is necessary that the student show his work for all the others, and the group will score his information representation and thought representation. The student should clearly understand what is his level of development regarding his ability to represent and also understand what he needs to improve. The final step is to check if someone obtained the maximum score in thought and information representation. If that does not happen, the applicant must present to the participants a production that shows the maximum score in both aspects and discuss with them the reasons for such a score. The program was built with the purpose of intervening in abstract formal operations, being not recommended for children, but young people and adults.

Until now, FLDP still does not have studies showing its effects. The current study is the first investigation about the effects of FLDP in fluid intelligence and the development of the levels of the inductive reasoning ability. Fluid intelligence was chosen because it is responsible for people´s ability to learn new things and reasoning in tasks that do not require a high load in previous knowledge. Fluid intelligence is also the ability that has the greatest association with general intelligence [16]. Besides that, inductive reasoning is one of the specific abilities of the fluid intelligence.

2. METHODS

2.1 Participants

This study used a convenience sample that does not represent all population of the young people (teenagers) that is enrolled at the Education of Young People and Adults' program (EYPA). The sample was composed by thirteen students, aged between 15 and 17 years old, from the EYPA, already litterate, enrolled in the module corresponding to Years 6 and 7 (Primary School). The corresponding acronym for EYPA in Portuguese is EJA (Educação de Jovens e Adultos) and it is a program for people that, for different reasons, did not complete primary or secondary school at the appropriate age. The referred EYPA center was administered by the local city council and was located close to the community called `Aglomerado da Serra´, in Belo Horizonte, Minas Gerais, Brazil. Despite the sample was formed by convenience, the case and control group was randomly chosen. The inclusion criterion was: (1) teenagers students that (2) did not complete primary or secondary school at the appropriate age, with (3) low socio-economic status. A 15 years old teenager did all the activities from unit 1 to 4 and concluded one activity from unity 5, almost completing the entire FLDP. This teenager is the case study and the other 12 teenagers are part of the control group, which had 50% (six) of female participants.

2.2 Instruments

2.1.1 Fluid intelligence tests kit – short version of the Higher-Order Cognitive Factors Kit (HOCFK)

The Fluid Intelligence Tests Kit is composed by three tests to measure the logical reasoning ability, general reasoning and the induction ability, and in a broader level, the fluid intelligence [17]. This Kit can be applied to primary students, from sixth grade on. The tests can be applied individually or in groups. The first test is the Inductive Test (I), composed by five items in the short version. Each item is composed by five groups of 4 letters. Among them there are four groups in which the letters are organized according to the same rule. The examinee has to identify the group of letters that does not follow the pattern and mark it with an X, within a time limit of 7 minutes.

The second test from the Kit is the General Reasonig Test (GR). Each one of the four items from the short version, is composed by a mathematical-logical problem statement and a space to solve it. The examinee should interpret the statement, solve the problem and choose one of the five multiple-choice answers. The time limit for the short version is 8 minutes.

The third test is the Logical Reasoning (LR) Test, composed by 6 items in the short version. Each item consists of a conclusion based on two abstract logical premises, with no empirical relationship to the real world. The respondent has to indicate if the logical conclusion is appropriate or inappropriate. The time limit is 9 minutes.

Gomes and Borges [17] showed evidence that the three tests from the Fluid Intelligence Tests Kit are unidimensional, and also have discriminant and convergent validity. Gomes [18] presents results that the tests from Fluid Intelligence Tests Kit are markers of fluid intelligence. The tests have Cronbach alpha above. 70 [17,18] and are capable to discriminate discrepancies in performance from fluid intelligence through Rasch person reliability index found [17].

2.1.2 Inductive Reasoning Developmental Test (IRDT)

The IRDT is an instrument that assesses seven stages of the development of the inductive reasoning in people from 2 year-old to 99 year-old or more [19,20]. The test application can be individual or collective. It is an extension in terms of complexity, of the Induction test that composes the Fluid Intelligence Tests Kit [17,18]. The domain of inductive reasoning domain was used because it is one of the best indicators of fluid intelligence [16]. The test is composed by seven levels of increasing complexity, indicating different developmental stages. Each level is measured by eight items, bringing to a total of 56 items.

Each IRDT item is composed of four letters, or sequence of letters, with a specific rule (correct items), plus one letter or sequence with a different rule (exception). The task is to discover which letter or sequence is the exception. The examinee has to identify the group that does not follow this pattern and mark it with an (X), within a 100 minutes time limit. The examiner starts the testing by giving the examinee eight items from the lowest level, which are the easiest items. If the examinee has more than two hits, which is a score greater than a hit by chance, the examiner asks the examinee to do the eight items from a higher level. If the examinee has more than two hits on this level, he should answer the items of the next level and so on.

Golino, Gomes, Commons and Miller [20] found in a sample of over 1,400 children, teenagers, adults and elderly that the IRDT presents a general factor and seven factors of first level, indicating the seven stages of the development of the inductive reasoning abilities (CFI=.96, RMSEA=.06). Using the dichotomous Rasch model, the researchers also found the seven stages of development by analyzing the presence of gaps between the seven groups of eight items. All items had adequate degree of fit to the Rasch dichotomous model with information weighted fit (InFit) ranging from .72 to 1.32 (M =.96, SD=.17). Unlike traditional models of Classical Psychometrics, models of Item Response Theory allow to check not only the model fit, but also the reliability for items as well as for people, representing a major breakthrough for the measurement of cognitive abilities [21]. The reliability of the 56 items was 1.00 and .81 for people.

The IRDT measures seven stages and the first three stages (singular representation, map representation and representation system) identify different levels of concrete reasoning ability. In these stages the person has to identify a pattern that is common to the group of letters, paying attention to the existing relationship in each group. The singular representation stage, for example, requires the respondent to be able to discriminate different letters (A A B A A) and to identify a pattern. During the map representation stage, the participant needs to identify the pattern of coordinated letters. During the system representation stage the participant has to identify the pattern of a system (group of maps) of letters (NOPR IJKM UVXY MNOQ QRSU). The next three following levels (singular abstraction, abstraction map and abstraction system) correspond to different levels of abstract reasoning stages. The group of letters has an abstract relationship and is no longer evident in the direct relationship of the letter, for example when comparing AB (sequence that does not have another letter between the first and second element) with EG (sequence that skips a letter, here F). The last stage is related with the participant metas systematic abilities, and it demands the participants to make a logical induction through the comparison of systems of abstract systems (for a review about the stages interpretation see: [22,23,24].

2.3 Procedures

The study applied a single case with control group design. The data was collated in the EYPA Center. As commented previously, the selection of the case and control group was random. While one unique participant composed the case, the control group was formed by 12 participants. The pretest was administered collectively by one psychologist. The application occurred during a meeting, which lasted approximately 50 minutes. Two classes of EYPA participated in this first moment. After the pretest, the researcher started the intervention, which had a total of approximately 60 hours, distributed along 4 months, two or three times per week and each session had a duration of 1 to 2 hours. Once the intervention finished the posttest was administered. Full protocols that characterizes the didactics and content of the program is encountered in the material for the applicant and the material for the student at the University of Minas Gerais, LAICO laboratory. Examples of the tasks can be founded in Gomes [25].

The tests used were printed and after the data was collated in an electronic spreadsheet, the paper tests were incinerated. This research was approved by the Ethics Committee from the Federal University of Minas Gerais and all the ethical procedures in relation to the participants were followed.

The data was analyzed using the single case methodology and the softwares Sigle Bayes ES and Singlims ES. This methodology has often been used in neuropsychology [26], to

estimate the rarity or abnormality of the individual´s score in a test is an essential part of the assessment process. The standard method to compare the scores of individual tests with a normative sample is to convert the raw score to a z score and evaluate it using a table of the area under the normal curve. When a normative sample used to compare the performance of an individual is large and representative, there is a tendency that the data is grouped in a normal curve, therefore the score of the individual can be positioned within the normal distribution. The individual z score is then obtained and his performance can be classified according to the level of rarity or abnormality. However when the normative sample used is small, it tends to be positively skewed, therefore standard deviation of the sample can be underestimated, as well as the rarity or abnormality of its performance. In such cases, instead of using the z score, the t distribution should be used to compare the samples.

The standard process considers a normative or control sample as if it were a population, therefore the average and standard deviation are used as if they were parameters instead of sample statistics. But there is a problem with this procedure: the distribution of the variance of the sample is positively skewed, which means that the z score and the rarity of the observation can be overestimated. When we are not able to assume a very stable estimation of the necessary parameters, a solution can be to use the standard deviation of a small sample as an estimation of the standard deviation of the population. That way we can calculate the probability using the t distribution instead of the z.

Sokal and Rohlf [27] described a t test modification for independent sample, that can be used to compare a single subject with the sample. Researchers use the t distribution (with n-1 degrees of freedom), instead of the standard normal distribution , in order to estimate the abnormality of the individual score and to test if it is significantly lower than the points obtained by the control sample. The formula used by Sokal and Rohlf [26] is:

$$t = \frac{x^p - \bar{x}}{\frac{d}{\sqrt{(n+1)-n}}} \quad (1)$$

where, for our purposes, x^p is the score of the single participant (case), $\bar{x}$, *d* and n are respectively: the average, the standard deviation and the size (number of participants) of the control group. The standard deviation is calculated using the denominator n-1, therefore adjusting to the small size of the sample.

The sample size of the control or normative group recruited for this comparison is usually smaller than 10 and often smaller than 5. In such cases, when the normative sample is small, the most appropriate method is to consider the subject as a sample of N=1 and to use the modified t test described by Sokal and Rohlf [27]. Crawford and Howell [26] recommend the use of the modified t test, instead of the traditional method which uses the z score, when the normative sample N is equal or smaller than 50. Although there is a postulate that the t test should be used for normal distribution, the authors cite studies in which the t test is surprisingly robust even when there is a moderate violation of the normality principal.

Crawford, Garthwaite and Porter [28], in order to stimulate single case studies, made available and described two softwares: The SigleBayes ES and Singlims ES. Such softwares are a convenient and reliable way to obtain the statistics of size effect.

3. RESULTS AND DISCUSSION

The fluid intelligence score was generated using the Rasch model and the participants answer to the items of the four tests used. The items of all tests are markers of the fluid intelligence. The participants data of the pretest (13 subjects) was added to the posttest data, bringing to a total of 26 raw scores considered in the analysis. The measurement of fluid intelligence had a Cronbach alpha of .74. The reliability index of separating people was .84, indicating that the measure of fluid intelligence was able to measure and discriminate the ability of each research participant, during both pretest and posttest. The reliability index of separation of items was .87, indicating that the items discriminate different levels of difficulty and ability.

The adjustment of the items and the pattern of the participants responses was assessed by InFit index. Values between .5 and 1.5 are considered adequate [29]. Regarding the adjustment of the items to the model, on average the items presented an InFit of .98 and a standard deviation of .34. Participants had a mean InFit of 1.01 and standard deviation of .33.The result indicates that, in general, both the items and the response patterns of the participants adjusted to the Rasch model.

Table 1 shows the InFit data of all the items used for measuring the fluid intelligence using the Rasch model. Some items showed a slightly improper adjustment, all those items are from the Logical Reasoning Test. The LR03, item from the logical reasoning test, showed an InFit of 1.51, the LR08 the InFit was 1.57, LR09 item presented InFit 1.57 and LR11 showed an InFit of 1.83, the most inappropriate of all. According to the literature, InFits above 2.0 can compromise the construction of Rasch measure, which is not the case any of these items. The items in which all participants gave the right answer or did not answer correctly did not have their InFit estimated. This occurred to items GR05, which everyone missed the item, and T-01, T-02, T-03, T-04, T-05, T-06, T-07, T-08, T-10, T-11 items, which everyone answered correctly.

Table 2 shows the InFit of the pattern of responses of the 13 participants for both pre and posttest. Regarding the participants´ score in the pretest, the participant 2 presented an InFit slightly inappropriate (1.59), as well as participant 10 (1.62). Meanwhile, during the posttest the participant 1 showed the worst fit (1.74), and participant 10 showed a slightly improper adjustment (1.54). It is important to highlight that the worst fit of the scores is the one obtained by the case of the research: the teenager who participated in the intervention. This poor fit is not significant, because it is not a value greater than 2.0, but it is an indication that possibly the intervention program has generated a pattern of responses not expected by the Rasch model. This condition would be expected in a cognitive program that can change the standard performance of a person.

The average performance of participants was -.18 logits and standard deviation of 1.27 logits, indicating a very varied performance. The average performance in the pretest was -.77 logits and standard deviation was 1.04. Meanwhile, the average performance in the posttest was .41, and a standard deviation of 1.29. It can be observed that the average performance of the posttest is greater than the pretest performance.

Table 1. Measure and Items adjustment (InFit) to the Rasch model

Items	Measure	Standard error	InFit	z	Items	Measure	Standard Error	InFit	Z
GR-01	-.16	.62	.64	-1.80	T-12	-4.08	1.07	.56	-.30
GR-02	-.40	.44	.72	-1.80	T-13	-4.08	1.07	1.13	.40
GR-03	1.92	.75	.78	-.40	T-14	-4.08	1.07	.56	-.30
GR-04	2.93	.81	.71	-.30	T-15	-2.71	.67	.77	-.40
GR-05	3.36	1.87		Max	T-16	-4.08	1.07	.56	-.30
GR-06	3.51	1.12	.51	-.50	T-17	1.97	.61	1.27	.80
LR-01	-1.75	.83	1.38	.90	T-18	.00	.44	.86	-.80
LR-02	-.67	.67	1.24	.80	T-19	-1.99	.55	.76	-.70
LR-03	-2.77	.84	1.51	1.00	T-20	-.40	.44	.87	-.70
LR-04	-2.16	.72	1.50	1.20	T-21	-.80	.45	.84	-.90
LR-05	-.89	.61	1.38	1.80	T-22	-1.71	.52	.91	-.20
LR-06	-.24	.64	1.05	.30	T-23	-1.00	.46	.94	-.20
LR-07	-.16	.62	1.13	.70	T-24	-.80	.45	1.19	1.10
LR-08	.16	.63	1.57	2.10	T-25	3.79	1.08	.54	-.40
LR-09	-1.27	.62	1.52	2.00	T-26	2.38	.68	.59	-.90
LR-10	-.24	.64	1.45	1.60	T-27	1.97	.61	.92	-.10
LR-11	2.56	.86	1.83	1.40	T-28	1.97	.61	1.16	.50
LR-12	-.16	.62	.95	-.10	T-29	2.93	.81	.71	-.30
T-01	-5.40	1.87		Min	T-30	1.97	.61	.76	-.50
T-02	-5.40	1.87		Min	T-31	1.63	.56	1.13	.50
T-03	-5.40	1.87		Min	T-32	2.93	.81	.74	-.30
T-04	-5.40	1.87		Min	IT-01	-.53	.60	.69	-1.70
T-05	-5.40	1.87		Min	IT-02	-1.75	.83	.91	.00
T-06	-5.40	1.87		Min	IT-03	-.53	.60	.95	-.20
T-07	-5.40	1.87		Min	IT-04	.23	.64	1.43	1.50
T-08	-5.40	1.87		Min	IT-05	1.41	.68	.62	-1.20
T-09	-1.99	.55	.89	-.20	IT-06	1.41	.68	1.18	.60
T-10	-5.40	1.87		Min	IT-07	1.97	.61	1.40	1.10
T-11	-5.40	1.87		Min	IT-08	3.79	1.08	.54	-.40

** GR: General Reasoning Test; LR:Logical Reasoning Test; T: Inductive Reasoning Developmental Test; IT: induction test*

Analyzing the performance of the case and the control group, both showed improvement in the posttest performance compared to the pretest. The performance of the case in the pretest was .22 logits, while the control group had an average performance of -.85 (SD= 1.04). In the posttest the case achieved a performance of 3.16 logits , and the control group had an average performance of .18 (SD=1.03). The case had a gain of 2.94 logits and the control group had an average gain of 1.03 logits (SD=1.04). The effect size was 1.84 (.88 to 2.76, t=1.76, p=.05 - one-tailed), when comparing the gain of the case in relation to the gain obtained by the control group. The population from the control group has 94.73% chance to get a lower gain compared to the gain obtained by the participant of the program. In other words, the population from the control group has 5.37% chance of getting a higher gain compared to the gain of the case. Thus, under similar conditions, and considering 100 interventions with these characteristics, it will be possible to find only five studies, in which the controls will have the average gain higher than the participant that performs FLDP. This inference about probability is only possible because the Bayesian method applied in conjunction with the frequentist approach of t test. Despite t test not being able to inform about the probability of the null hypothesis, Bayesian approach can do it. Because of the latter it is possible to say that the control group has 5.37% chance of getting a higher gain compared to the gain of the case.

Table 2. Measure and InFit of the People in relation to the Rasch Model

Participants	Condition	Test	Measure	Standard error	InFit	Z
1	case	Pretest	.22	.46	.58	-1.90
2	control	Pretest	-.81	.45	1.59	2.50
3	control	Pretest	-1.67	.48	.73	-1.10
4	control	Pretest	-.61	.45	1.36	1.60
5	control	Pretest	-1.23	.46	.76	-1.10
6	control	Pretest	.44	.47	.63	-1.50
7	control	Pretest	-.81	.45	1.08	.50
8	control	Pretest	-1.02	.46	.91	-.40
9	control	Pretest	-.81	.45	.62	-2.00
10	control	Pretest	-3.36	.60	1.62	1.60
11	control	Pretest	.66	.48	.99	.10
12	control	Pretest	-1.02	.46	.86	-.60
13	control	Pretest	.01	.46	.81	-.80
1	case	Posttest	3.16	.52	1.74	2.50
2	control	Posttest	.45	.47	1.16	.70
3	control	Posttest	-1.36	.50	.99	.00
4	control	Posttest	.01	.47	1.18	.70
5	control	Posttest	-.43	.47	.65	-1.50
6	control	Posttest	1.97	.47	.81	-.90
7	control	Posttest	1.11	.47	.97	.00
8	control	Posttest	.01	.47	.97	.00
9	control	Posttest	.23	.47	1.36	1.30
10	control	Posttest	-1.62	.51	1.54	1.80
11	control	Posttest	.89	.47	1.08	.40
12	control	Posttest	-.21	.47	.70	-1.20
13	control	Posttest	1.11	.47	.68	-1.30

In addition to the gain related to fluid intelligence, another important information is whether FLDP was able to generate qualitative changes (like-stage) in the development of the participant´s inductive reasoning and if this change is greater than the changes observed in the control group. The score of the cognitive stage of inductive reasoning of each participant corresponds to the stage of the highest level that the hits of the participant surpass the chance level. Therefore, when the participant gets a hit equal or greater than three items of a grouping of eight items, his performance is considered satisfactory at this stage. If this is the most advanced stage which the participant has more than two hits, then the participant's score is determined by that stage. The scores obtained by the participants were: a score 2, which characterizes the stage of map representation; the score 3, featuring the stage of the system of representation; and score 4, featuring the stage of singular abstraction in inductive reasoning. Table 3 presents the scores of each participant in the pretest and posttest.

Of the 13 participants, only two of them got score 2 in the pretest and the remaining obtained a score 3. The score 2 indicates a stage called map of representation. Individuals with this score showed the ability to map logical concrete patterns. The score 3 indicates the stage of system representation. Individuals at this score show the ability to build logical concrete systems, for example, they can arrange in a flexible way, two maps of logical pattern into a system. All the participants remained in the stage of development that they were during the pretest, except the one who participated in the FLDP. This participant obtained a score 3 during pretest, but scored 4 during posttest, changing from the system representation stage to the singular abstraction stage. The case was the only participant of the study who qualitatively altered the ability of inductive reasoning. She showed, in the posttest, the ability

to build unique abstract units, through the coordination of two systems of representation. The gain of the control group was zero in relation to the stages, with a standard deviation of .38 (stage), while the case showed a gain of one stage. The effect size of the difference in gain between the case and control group was 2.63 (1.40 to 3.84, 95% confidence level, one-tailed p = .01) and the percentage of population control to obtain a larger gain that the case is 1.40%.

Table 3. Scores of the cognitive stages of Inductive Reasoning

Participant	Condition	Pretest	Posttest	Increase
1	Case	3	4	1
2	Control	3	3	0
3	Control	2	2	0
4	Control	3	3	0
5	Control	3	3	0
6	Control	3	3	0
7	Control	3	3	0
8	Control	3	3	0
9	Control	3	3	0
10	Control	2	2	0
11	Control	3	3	0
12	Control	3	3	0
13	Control	3	3	0

4. CONCLUSION

The results of this study indicated favorable first evidence that the FLDP increases fluid intelligence. If the experimental condition of this study is replicated 100 times, it is likely that the control group achieves an average gain greater than the FLDP participant, in only five of these replications.

The results also indicated that FLDP qualitatively alters the level of inductive reasoning, being able to provide a shift from concrete thinking skills in inductive reasoning system of representation to the abstract thinking (singular abstraction). The participant who took part in the FLDP changed one stage of cognitive development in the inductive reasoning ability, while no participant in the control group changed their cognitive stage. Under the same conditions, if this study is replicated 100 times, it is likely that only one control group will obtain a stage change equal or greater than the participant of the FLDP.

These results are in agreement with the meta-analysis review of Romeny and Samuels [8] which found moderate effect sizes of four fluid abilities such as planning (d=1.04 p<.01), spatial-perceptual ability (d=.58, p<.01) and reasoning skills (d=.61, p<.01). However, other fluid skills that were the focus of the FLDP showed a weak effect size in the meta-analysis conducted by Romeny and Samuels [8], such as metacognition (d=.49, p<.01) and general cognitive ability (d=.49, p<.01). The present study did not use specific measures of fluid abilities, imposing a limit on the comparison between the meta-analysis results [8] and our findings. In the same direction, the results are in agreement with the findings of Kozulin et al. [10] showed effect sizes ranging from .11 to .46 (p <.05) for some WISC's fluid intelligence subtests. Buschkuehl and Jaeggi [6], using review studies, proposed some criteria to consider cognitive intervention successful. There are parameters to evaluate the scope and

limitations of intervention studies, and some of them are: (1) Significant sample size; (2) Randomization of the participants when distributing the research groups, in order to minimize the study problems of internal validity; (3) Active control group, which is involved in activities during the training period, but such activities should be different from the ones applied to the experimental group; (4) Careful selection of multiple intelligence tests for the pretest and posttest, because when a single measure is used, the effects of the intervention are limited to gains in test scores; (5) Assessment of the long-term effects, not only in the experimental group; (6) Use of complex working memory tasks, when the target of the study is this; and, finally (7) Evaluation of long-range transfer. Short-range transfer refers to the increase in performance very similar to the training task. Therefore, the long-range refers to the increase in performance of tasks of different nature.

Considering these parameters, it is possible to conclude that the intervention described in this study met the criteria for an active control group, as the control group engaged in intellectual activities (all the participants were enrolled in the EYPA program) and evaluation of results through multiple instruments. Regarding the sample size, the single case method was used in this study, following all necessary scientific rigor, using reduced samples without compromising the reliability of the results. For future research, it is suggested to review the long-term effects of FLDP in experimental and control groups, and evaluation of the long-range transfer and generalization of cognitive gains for different contexts.

We conclude that FLDP offers evidence to be a useful tool for the modification of intelligence of young adult Brazilian students. The use of single-case methodology with control group allowed to build initial evidence about the effects of the program, both in fluid intelligence and in the development of cognitive stages in the ability of inductive reasoning. Further studies are needed, incorporating other research designs, as well as new variables to analyze the effects of the program. Finally, we emphasize that the results provided by this study represent an important contribution to the field of research about the possibilities of cognitive intervention, since it showed positive and promising results on the effects of an intervention on the mental performance of the participant, providing results important to further this field of study. Furthermore, it is expected that such initiatives are implemented in the Brazilian educational context, which deals with the challenge of increasing the cognitive abilities of their students

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Colello SMG. Reforma Curricular Brasileira: Whither Teacher Training? International Studies on Law and Education. Portuguese. 1999;55-56.
2. Gomes CMA. Betting on the development of inteligênciaIn search of a new educational curriculum for the development of human thought. Rio de Janeiro: Editora Oil Lamp English; 2007.
3. INAF Brasil. Relatório. Website of the “Instituto Paulo Montenegro”. Accessed 04 April; 2014.
Available:http://www.ipm.org.br/ipmb_pagina.php?mpg=4.02.01.00.00&ver=por&ver=por

4. Jensen AR. Raising the IQ: The Ramey and Haskins Study. Intelligence. 1981;5:29-40.
5. Strenze T. Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. Intelligence. 2007;35:401-426.
6. Buschkuehl, M, Jaeggi SM. Improving intelligence: A literature review. Swiss Medical Weekly. 2010;140:266–272.
7. Nisbett RE. Intelligence and How To Get It: Why Schools and Cultures Count. New York: W. W. Norton & Company; 2009.
8. Romney DM. Samuels MT. A meta-analytic evaluation of Feuerstein's Instrumental Enrichment program. Educational and Child Psychology. 2010;18(1):19-34.
9. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New Jersey: Lawrence Erlbaum Associates; 1988.
10. Kozulin A, Lebeer A, Madella-Noja A, Gonzalez F, Jeffrey I, Rosenthal N, et al. Cognitive modifiability of children with developmental disabilities: A multicentre study using Feuerstein's Instrumental Enrichment—Basic program. Research in Developmental Disabilities. 2010;31:551–559.
11. Kozulin A. Cognitive Modifiability of New Immigrant Adults. Mind, Culture, and Activity. 2009;16:117–129.
12. Lizarraga M, Ugarte M, Cardelle-Elawar M, Iriarte M, Baquedano M. Enhancement of self-regulation, assertiveness, and empathy. Learning and Instruction. 2003;13:423-439.
13. Shiell JL. A meta-analysis of Feuerstein Instrumental Enrichment. Doctor Thesis, Philosophy Departament, University of British Columbia; 2002.
14. Hernstein RJ, Murray C. The Bell Curve: Intelligence and Class Structure in American Life. New York: Free Press; 1994.
15. Gomes CMA. The Program for the Development of Formal Logic (PROLOF) and university education. In: Proceedings of the International Conference on Special Education. International Conference on Special Education. Sabara; 2003. (Portuguese)
16. Carroll JB. Human cognitive abilities: A survey of factor-analytical studies. New York: Cambridge University Press; 1993.
17. Gomes CMA, Borges O. Psychometric qualities of the set of fluid intelligence. Avaliação Psicológica. (Portuguese). 2009;8(1):17-32.
18. Gomes CMA. Factorial structure of the factors Cognitive Battery High-Order (BaFaCaLo). Avaliação Psicológica. (Portuguese). 2010;9(3):449-459.
19. Gomes CMA, Golino HF. Exploratory Study on the Development Test of Inductive Reasoning (TDRI). In D. Colinvaux. Proceedings of the VII Brazilian Congress of Psychology of Development: Development and Rights Humananos. Rio de Janeiro; 2009. Available: http://www.abpd.psc.br/files/congressosAnteriores/AnaisVIICBPD.pdf.
20. Golino HF, Gomes CMA, Commons MLC, Miller PM (in press). The Construction and Validation of a Developmental Test for Stage Identification: Two Exploratory Studies. Journal of Adult Development. (In Press).
Available: http://adultdevelopment.org/jad_special_issue.php.
21. Andrich D. Controversy and the Rasch model: A paradigm of incomparable paradigms. Medical Care. 2004;42(1):17-116.
22. Fischer KW. A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review. 1980;87:477-531.
23. Fischer KW, Bidell TR. Dynamic development of action, thought, and emotion. In: Damon W, Lerner RM, editors. Theoretical models of human development. Handbook of child psychology. 6th ed. New York: Wiley; 2006.

24. Fischer KW, Yan Z. The development of dynamic skill theory. In: Lickliter R, Lewkowicz D, editors. Conceptions of development: Lessons from the laboratory. Hove, U.K.: Psychology Press; 2002.
25. Gomes CMA. The development program of formal logic: concepts and architecture. Education in Focus. (Portuguese) 2005;8:13-20.
26. Crawford JR, Howell DC. Comparing an Individual's test score against Norms derived from Small Samples. The Clinical Neuropsychologist. 1998;12(4):482-486. DOI: 10.1076/clin.12.4.482.7241.
27. Sokal RR, Rohlf JF. Biometry. San Francisco, CA: W.H. Freeman; 1995.
28. Crawford JR, Garthwaite PH, Porter S. Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. Cognitive Neuropsychology. 2010;27(3):245-260.
29. Wright BD, Linacre JM. Reasonable mean-square fit values. Rasch Measurement Transactions. 1994;8(3):370.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=488&id=21&aid=4724