



A Reflection on Teaching Practices in Middle Schools in Disadvantaged Areas - With a Focus on the Level of the Cohort Compared to That of the Single Class

Emmanuèle Auriac-Slusarczyk

Associate professor

Laboratoire Activité, Connaissance, Transmission
Education (Acté, EA 4281), University of Clermont Auvergne

Hélène Maire

Associate professor

Laboratoire Lorrain de Psychologie et Neurosciences de la
dynamique des comportements (2LPN, UR 7489), University of Lorraine

Inès Barroche

4th year speech therapy student

School for Speech Therapy, Clermont Ferrand
Laboratoire Activité, Connaissance, Transmission
Education (Acté, EA 4281), University of Clermont Auvergne

Lison Haptel

4th year speech therapy student

School for Speech Therapy, Clermont Ferrand
Laboratoire Activité, Connaissance, Transmission
Education (Acté, EA 4281), University of Clermont Auvergne

Tiphaine Hudry-Critoux

School monitor, Académie de Clermont Ferrand
Laboratoire Activité, Connaissance, Transmission
Education (Acté, EA 4281), University of Clermont Auvergne

Bernard Slusarczyk

Middle school principal, Académie de Clermont Ferrand
Laboratoire Activité, Connaissance, Transmission
Education (Acté, EA 4281), University of Clermont Auvergne

ABSTRACT

This article addresses the question of the adaptation in a teacher's pedagogical practices of sixth-grade pupils in priority education zones. This research approach, within the framework of a research intervention program, sought to verify the correlation between pupils' cognitive, language skills, and their ability to benefit from different philosophical teaching practices. Two types of measures are deployed in relation to sixth-grade pupils. Two standardized tests (language:

linguistic/comprehension - EVAC and ECL-2 tests - measure of global fluid intelligence - Raven's matrix test) evaluate the scores. This assessment is cross-referenced with the measurement of the pupils' academic skills in language and mathematics assigned by teachers. Our results are based on a full sample of 124 pupils in seven classes, sometimes reduced to 90 pupils by removing pupils from linguistic minorities. The results indicate, by comparing two samples, significant differences in the levels of the pupils in priority education zones compared to standard scores: a minimum gap of 20 percentiles is noted for each subtest, which implies a form of delay in the academic performance of the pupils in priority education zones. Compared to the general average of French sixth-grade pupils, a lack of lexical skills is particularly apparent, indicating a form of abandonment. A differentiation is apparent between the different classes tested. Therefore, the scale of the class or schooling context must therefore be distinguished and studied separately. We discuss the desirability of identifying more precise pupil profiles, to enable teachers to move beyond an overly global view of the class.

Keywords: disadvantaged areas; middle school; academic delay; linguistic skills assessment; fluid intelligence.

INTRODUCTION

It is well-established that introducing philosophically adapted teaching approaches at school leads is beneficial. Gorard and colleagues (2017) revealed that practicing Philosophy for Children (P4C) had positive effects on academic skills such as reading and mathematics, and this gain was *larger for the most disadvantaged pupils* (House, 2016; Siddiqui et al., 2019). As sometimes described, professors conduct teaching activities through the repartition of lower, middle and higher-level pupils, particularly in regard to verbal exchanges, which circularly increases pupils' level status and reputation in the set of each class (Clanet, 2005). Interestingly, the necessity to introduce philosophical workshops for middle school teachers in Middle School has led researchers to describing pupils more accurately through their competencies. How can teachers introduce a philosophical approach at school without disqualifying innovative practices (Fournel et al., 2020)? The work initiated by Matthew Lipman is well-entrenched in the French educational environment (Leleux, 2005, 2014; Lévine 2005; Lipman, 2003; Lipman et al., 1980) and the positive benefits associated with P4C are widely documented (e.g., Colom et al., 2014; Gregory et al., 2016; Millett & Tapper, 2012; Topping & Trickey, 2007, 2014; Trickey & Topping, 2004; Yan et al., 2018), as well as other positive effects of discussion processes (Murphy et al., 2009): but in which manner could disadvantaged public schools be good candidates to benefit from such philosophical programs?

In the French context, previous studies, conducted in preschool, primary and middle schools, have shown that the measures of the impact on cognitive integrity and moral development take into account a recurring problem (Auriac, 2006; Auriac-Slusarczyk et al., 2011; Slusarczyk et al., 2015). As demonstrated in the field of argumentation, the factor general intelligence (the g factor) is confused with linguistics and verbal competencies (Auriac, 2002; Neuman, 2003): it is therefore necessary to provide a thorough description of this circularity and regularly inform the educational community about it. Another previous study-observed that the comprehension

of conventional metaphors (Rey et al., 2007¹) in sixth-grade pupils in disadvantaged areas was lower than standard levels, with as much as a two-year discrepancy (Slusarczyk et al., 2020). Considering these preliminary results, the main goal of this paper is to study and to discuss more adequately the students being taught in disadvantaged areas², and also to determine the level of the risk factors of failure brought about by the lag in the linguistic abilities of immigrant children (Plumelle, 2016). So, this article aims to develop a novel description based on standardized tests, assessing cognitive and linguistic skills in order to develop intelligence and implement quality adaptation to the teaching approaches/practices at school.

THEORETICAL FRAMEWORK

Characterizing pupil profiles calls for the evaluation of both the linguistic and the cognitive skills, using standardized tests, which in turn would allow the teachers to continue their commitment to teaching in an independent manner, all the while taking into consideration the scientific description of their pupils. As sometimes considered, an entertaining ambiguity in educational research consists of creating pedagogical controversy (Fedi, 2012, e.g., Piaget vs. Lipman), which increases the difficulty of implementing certain professional choices, particularly in transversal activities or innovation (Fiema et al, 2020). Thus, the question is, is it more important and congruent to consider the sixth-grade cohort or a single class in order to associate good professional posture and peer effect classes?

In the French school system, the professional skills (Clot, 2007) mastered by teachers are generally conducted to be adapted to the context of teaching in a singular class. But, in disadvantaged areas, teachers also adhere to the idea of adjusting the relationship with knowledge, albeit sometimes too radically or too simply, with the sociological adjustments between teachers and pupils (Charlot, 1999; Charlot et al., 1992; Lahire, 1993; Van Zanten & Fijalkow, 1994). We develop the following a) the necessity to fix a horizon of adaptive programs to the adolescence period through specific levels of cognitive competencies relevant to the age of the pupil, b) the necessity to benefit from standardized tests to describe a school population better, c) the necessity to qualify and measure the gap between standards/norms and instruction in disadvantaged areas, to promote well-adapted teachers. The literature has already recorded the impact of an adapted philosophy on general intelligence (the factor g, Mortier, 2005). Gender impact is sometimes taken into account (Abaspour et al., 2015). It is clear that the balance to be found between intellectual maturity at adolescence, coupled with the linguistic underperformance proven in schools in disadvantaged areas, results in ad hoc adaptations of pedagogy.

The adaptation of philosophical practices at middle school: linguistic or intelligence problem?

Studies show that applying a philosophical approach to teaching practices the practice of philosophy has many developmental benefits at all grade levels and in many areas. For example, in preschoolers, fostering verbal reasoning skills (i.e., ability to give a reason for an opinion: justification, comparison, analogy, causal connection or the wording “because of that”, Säre et

¹ *Cultural* or *conventional* metaphors indicate expressions used in language that do not require semantic interpretation, but rather decoding in the context of the situation, such as “to not be the full shilling” or “to act like the cock of the walk”.

² The disadvantaged areas are designated in France through the abbreviation “REP+”, meaning **R**éseaux d'**E**ducation **P**rioritaire renforcés, that refers to reinforced (+) priority (P) education (E) networks (R).

al., 2016) is effective. In primary school, promoting subjective argumentation skills in children 7-8 years old (Walker et al., 2013) and increasing the participation of pupils in classroom dialogue (Topping & Trickey, 2014) facilitate critical thinking development (Lam, 2012) and moral understanding (Bleazby, 2020). Furthermore, philosophical training has also been shown to improve non-cognitive, socio-emotional skills such as emotion comprehension and social competence in preschool children (Gimenez-Dasi et al., 2013). Additionally, academic self-esteem anxiety seems to be reduced by P4C, whereas higher self-esteem, self-efficacy and adaptive behaviors are promoted (Sasseville, 1994; Trickey & Topping, 2004, 2006). Philosophical programs also reinforce adherence to rules and self-regulation in teenagers aged 14 to 16 years (Heron & Cassidy, 2018). However, various factors of improvement and development are simultaneously studied, and their respective effects may be intertwined. It is to be determined which of these correspond specifically to cognitive skills that enable age-appropriate development in adolescence. A long-term study including participants from primary school to university revealed that philosophically adapted programs addressed to children/teenagers were more prone to prosocial behavior (Colom et al., 2014). Taken together, the results suggest that philosophical programs may foster academic achievement, not only in a direct manner (e.g., Jenkins & Lyle, 2010; Millet & Tapper, 2012), but also through some transversal skills, underlying academic performance. It may be likely that these putative indirect effects of philosophical programs have a larger impact on disadvantaged pupils, since they have been reported to be more receptive to this practice (see House, 2016; Gorard et al., 2017; Siddiqui et al., 2019).

Philosophical programs are not easy to implement for teachers. Teachers often consider the pupils' progression through explicit verbalizations, and do not spontaneously admit that abstaining from assessing³ (Agostini, 2012) is also a good and alternative way of cognitive development. Thinking or speaking? What is better? The role of linguistic skills is at the heart of their professional preoccupations. Linguistic correction is, also, for professors, a major problem in disadvantaged pupils placed under their supervision. A previous exploratory experiment, carried out in middle school, verifies the standardization of the population, by increasing the homogeneity of cognitive integrity in the classes where a philosophical approach is considered compared to control groups (Slusarczyk et al., 2015). Is the standardization of levels a good issue?

How and why to test pupils in disadvantaged areas?

As often, the data collected in the field of education are contextualized and cited in particular practice or characteristic areas. The collection of a larger sample of the pupils' original and ecological situations is necessary to conduct quantitative studies in order to estimate the peer-effect of pupils across various scales.

In fact, disadvantaged areas gather a public with academic disadvantages and the problems of racial differences (Tudge & Doucet, 2004), which could drive academic misidentifications (Osborn, 1997, for example). Self-esteem, in the case of racial minorities in a school community, could decrease academic performance. Further, a familiar context of low economic status could interfere with academic performance, and this, from a very young age (kindergarten), when

³ This was in reference to Philosophie and Philosophy of *Montaigne* and sceptic orientation in Philosophy, represented by *Nietzsche*.

comparing the socio-economic backgrounds (Miller & Votruba-Drzal, 2017). Comparatively, during kindergarten, in middle school, differences in socio-economic status presents an even bigger problem (see Duncan & Magnuson, 2011; Van Ewijk & Sleegers, 2010, particularly when a pupil arrives during the school year, as it is often the case in disadvantaged area, a focus in this research project (see our methodology). As noted by Van Ewijk and Sleegers (2010, p.135), the meta-analysis, conducted on thirty studies of pupils aged from 8 to 18, estimate different amounts of effect, running from no effect to a small or large effect depending on the scientific methodology or sample considered (Van Ewijk & Sleegers, 2010, p. 135). In those studies, twenty-one percent estimated the impact of socio-economic status on scholarly achievement in reference to tests in mathematics and language. The researchers emphasize the need to properly estimate the respective effects at the cohort (more than 100 pupils) and/or class (less than 40) level, and to compare any differences. In addition, the variations between racial negative perceptions, low socio-economic status and early low academic performance could interfere with, and have a negative influence on, effective instruction in the classrooms (Creemers & Reezigt, 1996; Sorhagen, 2013). Therefore, it is important to consider a sociological view (Dreeben & Barr, 1988; Bourdieu, 1979, 1989) in order to consider teaching practices that combine several influencing factors. In their meta-analysis, Van Ewijk and Sleegers demonstrated, as was done in the francophone system (Chauprade et al., 2015), that ideological features could explain the considerable difference, also bias, between estimated effects, which varied from 0.03 to 0.59 standard deviation, covariate with country and methodological choices. In many studies, average socio-economic status (SES) was measured at the level of the cohort or school (see Van Ewijk & Sleegers, 2010, p. 139).

Alternatively, in order to reveal the various conditions to achieve singular performance in school program, an adapted methodology of standardization testing is necessary (Dunn et al., 1986). Often, the measurement of language and mathematical skills during the preschool period (LeFebvre et al., 2009), for example, or in middle school (Van Ewijk & Sleegers, 2010), is the only one taken into consideration, which conditions educational research. The specific disadvantaged areas in middle classes focus only on language measures: so, in adolescence the poor academic performances measured concern low academic language abilities due to ethnic composition and repartition. So, applying measures in middle classes in disadvantaged areas leads to questioning the possibility of controlling the influencing factors in the pursuit of the school program, as is commonly done (Hughes et al., 2018; Magnuson et al., 2007, for example). Rather than introducing academic factors as dependent variables, using each pupil's actual academic level as an independent variable would ensure a better fit with sociological cautions and orientations. In this field, French studies have demonstrated the possibility of adapting for an allophone population by overcoming the linguistic disabilities in choosing reflexive drawing (Cayet, 2019, 2020, 2021) which contrasts with studies reintroducing linguistic written competencies through cultural art design in more ordinary middle schools (Auriac-Slusarczyk et al., 2019, 2018a, 2018b; Maire et al., 2018, 2019).

In a word, in previous studies, various factors such as race, social-economic status, language disability (with possible low, early performance as a singular trajectory, see Helbling et al., 2019, for example), are often confounded. Moreover, measured with a single test method, they also become a limitation, generalizing too quickly from a single community located in a single geographic area (see Herker & Tymms, 2004).

Academic measures and standardized intelligence tests

Beyond the academic abilities assessed exclusively with academic measures, a large number of studies indicate that general intelligence is one of the best predictors of educational achievement, showing moderate to strong correlations between IQ scores and school performance grades (Bartels et al., 2002; Neisser et al., 1996; Sternberg et al., 2001), even with variables such as prior school performance, sex, and SES controlled (e.g., Deary et al., 2007; Harackiewicz et al., 2002; Kuncel et al., 2004). For instance, in Deary et al.'s study (2007), the results of the intelligence tests (i.e., Spearman's g) taken at the age of 11, correlated 0.81 with national school examinations taken at the age of 16. However, although direct associations between children's intelligence and academic performance are well established, additional variables such as cognitive stimulation at home have been found to modulate these associations, especially in children living in poverty (e.g., Ayoub et al., 2009). Moreover, the association between intelligence and academic performance may also depend on which academic domain is tested. For instance, the latent intelligence trait in Deary et al.'s study (2007) explains a higher performance in mathematics (58.9% of the variance) and linguistics (in English, 48.3%) than in art and design (18.2%). Taken together, these results suggest that a deeper comprehension of the abilities in children in disadvantaged areas could be provided by crossing results from intelligence measures with academic results in various specific fields.

METHOD

We first present the standardized testing tools before describing the participants and the scoring system. Methodological precautions are presented in relation to the ecological and contextual setting of our study.

Participants

The seven classes (anonymously designated by the letters *A, B, C, D, E, F, G*), encompassing all of the sixth-grade pupils in the same middle school, sat several tests that assessed the level of their linguistic skills (lexical, grammatical), comprehension skills (stories), and intelligence (g factor). Each teacher used tests to better adapt further philosophical programs.

In a first test, the 143 pupils enrolled in the seven participating classes was removed from the total process for various reasons⁴ 124 of the subjects tested furnished the items of valid data (see Table 1) used to analyze the linguistics and comprehension scores. Within this cohort, 34 pupils were of foreign nationality: *Algeria, Kosovo, Bulgaria, Romania, Turkey, Russia, Mali, Spain, Portugal, Armenia, Guinea, Morocco and Italy*. Therefore, we looked over the scores by removing pupils of foreign nationality to check whether the general trends in the remaining data used impacted the results. Across the seven classes, the gender balance of the pupils was divided as follows: 48,4% girls and 51,6% boys (see Table 1). To test the Raven- g factor, 92 subjects were tested, with a reduced gender balance.

Table 1. Distribution of 124 first and 92 second collecting data pupil: gender (girls, boys) according to sixth grade class (*A, B, C, D, E, F* and *G* for information).

⁴ Several pupils who did not sit the EVAC test were split across several classes: *F, C, A, E* and *D*. Some pupils for whom the school did not have evaluation levels for French and mathematics. One did not understand the instructions for the story ECL-2 comprehension test, and six pupils did not sit the comprehension test.

	Class							Total
	G	A	B	C	D	E	F	
GenderGirls	9/0	8/7	9/9	8/4	10/7	7/7	9/8	60/42
Boys	5/0	11/11	11/11	10/7	8/8	12/12	7/3	64/50
Total	14/0	19/18	20/18	18/11	18/15	19/16	11/11	124/92

Because class G did not participate in the collective second Raven-g Test, we have chosen to present a statistical treatment of the data scores of pupils of the *A, B, C, D, E and F* sixth-grade classes (see our § Result III. 3.) In fact, only 63 pupils were tested in all standardized tests: EVAC, ECL-2 and Raven-g factor. Each pupil is sited at an academic level in language and mathematics.

Material and procedure

The researchers evaluated the linguistic and comprehension skills of all the sixth-grade pupils considered as a disadvantaged pupil cohort. Several subtests taken from the EVAC⁵ scale (Flessas & Lussier, 2003) and one of the subtests highlighting comprehension skills taken from a ECL-2 battery of tests⁶ (Khomsy et al., 2005) were selected and used. To complete, general fluid intelligence⁷ among pupils was evaluated by score means of a Raven *g*-factor measurement test (Raven's Progressive Matrices, 1998) taking the opportunity to cut the score into intermediate sub-scores (Chartier et al., 2014), regarding the performances of series *A* (easier) to *E* (more difficult). In each case, the objective was to obtain a good profile of the sixth-grade cohort, and to acquire a good description of the repartition of levels (low, middle and high good) in each class. Achievement levels in French and mathematics, recorded in end-of-term reports at the start of the school year, corresponded to other variables to verify correlations with the standardized test set. All testing sets were conducted in a collective manner. The EVAC and ECL-2 were conducted in each class during school time, replacing the language lesson, and the Raven-g Test was conducted in the school's auditorium, gathering all the classes of the sixth-grade cohorts.

Modality of testing, statistical tools, validity of our ecological experimental protocol

To determine the performance levels, a) the linguistic aptitude (EVAC) test is described subtest by subtest (*Riddles, Missing Words, Expressions, Linking words, Lexical skills*⁸) and with total score, b) the subtest for story comprehension (ECL-2) gives a global score. We used Reference Tables for elementary and middle school level for the EVAC to compare the distribution of scores with a normal distribution. For the ECL-2, we used a personalized scoring system by awarding one point (1) if the answer reflected precise understanding of the expression, no points (0) if the literal or plausible answer did not consider the interpretive subtlety of the suggested expression and took away a point (-) if the answer chosen was entirely incorrect. For Raven-g Test, the standard Referent Tables are used to determine the gap with average means and standard deviation. For mathematic and language academic assessment, a continuous scale from 0 to 20 is considered. Descriptive and inferential statistics were carried out, using SPSS

⁵ Verbal Test of Cognitive Aptitude.

⁶ Evaluation of Linguistic Skills, written for middle school. ECL-2

⁷ *Fluid* intelligence refers to the ability to reason and see connections independently of prior knowledge and experiences, while *crystallized* intelligence is based on knowledge and experiences.

⁸ The content of these subtests was presented with the results in order to facilitate reading.

software version 23, according to our perspectives of description and needs of extraction of significant tendencies. All scores of each test are considered, and we adapted each statistical test (Student *t*, correlation, regression analysis) as the mathematical bases (see Brauer, 2002; Cohen et al., 2003).

The ecological context characterizing the modality of collecting our data enabled the emergence of the evident benefits as well as the evident limitations. As stated above, a gender balance benefit to boys could be an important consideration in other, future studies. In the frame of our study, this tendency has no strong importance. The study could be reproduced. To replicate the study, researchers would have to consider a whole sample of pupils in a comparably disadvantaged area. The number of pupils tested is rather high compared to other studies. Generally, a cohort is considered as a group of 100 pupils, whereas a class is composed of 40 pupils. (see Van Ewijk & Sleegers, p. 139). Our study presents some original and ecological features having considered other quantitative and qualitative aspects. In particular, the aspect of standard deviation (see Chartier, 2014) and the composition of the sample are important when deciding on any adaptations to the teaching approach, with a view to the specificity to a cohort or to a singular class.

RESULTS

Our results focus on three scales: individual vs. class vs. cohort. They attempt to highlight what is interesting about crossing measures: academic vs. test-measured, and subtests, to achieve finesse in describing pupil profiles. Finally, any gender was not evaluated.

Descriptive statistics for the seven sixth-grades (124 subjects) indicate that the academic level is, on average, the same in French and mathematics, with a correlation ($r(124) = .653, p < .0001$). The sample mean is 6.69 vs 6.64 for French and mathematics, respectively, and spans a minimum-maximum range of comparable means: 3.30 to 9.50 vs 3.50 to 9.80 for French and mathematics with a standard deviation of 1.4 vs 3 and a variance of 1.97 vs 1.73. There is a very slight standardization (lower variance) in favor of mathematics. For the 124 valid data items, and considering only the total score for each of the tests (EVAC and ECL-2), the correlation between the two tests is significant ($r = .487, p < .0001$). The correlation between overall scores (EVAC and ECL-2) and grade levels is also significant. The French level is significantly correlated with language ability measures ($r = .597, p < .0001$), as is comprehension ($r = .313, p < .0001$); the math level is significantly correlated with language ability measures ($r = .590, p < .0001$), as is the comprehension ability ($r = .273, p < .002$). Regarding the comparison between classes, as global and preliminary trends, T-tests considering the class as an independent sample show that one class (G) differs, notably on the equality of variance test for mathematics and French. In addition, the EVAC and ECL-2 test scores differ significantly between each class (see peer effect classes, part III.4.). The T-tests do not identify a gender impact in the overall scores. For example, the T-tests for the reduced sample of 90 correspond to French level: $t(88.4) = -1.236, ns$; math level; $t(88.4) = -.471, ns$; EVAC score $t(88.4) = -.231, ns$; ECL-2 score $t(88.4) = .961, ns$. No effect was found on several EVAC subtests.

Our results which are presented take into account the finesse of each of the tests and subtests that make up the interest of our methodology. Parts III.1. and III.2. show the first trends, comparing the results from the full sample of 124 pupils, and its reduction to 90 pupils, removing the minority pupils as it cannot be ascertained that they have attended school enough

to allow them to master the language. Parts III.3. and III.4. deliberately detail, step by step, the salient facts that will serve the discussion.

Correlations between subtests assessing linguistic aptitudes, comprehension and academic levels

There is a significant correlation ($p < .0001$) between all the subtests of linguistic aptitude, which are correlated with academic levels in French, mathematics and in comprehension tests (Table 2).

Table 2. Correlation between each of the linguistic sub-skills tested, comprehension score and academic levels (N=124)

	Riddle	Missing words	Expressions	Linking words	Lexical skills	EVAC Total score	Comprehension Total	Academic level in French	Academic level in mathematics
EVAC Riddles	1	.564**	.520**	.584**	.637**	.816**	.527**	.521**	.432**
EVAC Missing words		1	.531**	.607**	.550**	.789**	.326**	.491**	.337**
EVAC Expressions			1	.534**	.546**	.716**	.329**	.423**	.409**
EVAC Linking words				1	.673**	.838**	.341**	.490**	.543**
EVAC Lexical skills					1	.880**	.431**	.497**	.612**
EVAC Total score						1	.487**	.597**	.590**
Comprehension Total score							1	.313**	.273**
Academic level in French								1	.653**
Academic level in mathematics									1

Note. ** Correlation is significant at 0.01 (bilateral)

The values of the correlations (proportion of each subtest 42% to 59% for French and 33% to 61% for mathematics) support the fact that the level in French or in mathematics reaches an equal proportion (.597 vs. .590). Some subtests have a stronger relation to aptitude in mathematics or logic, rather than the two subtests “*linking words*”/“*lexical skills*” assessing memory skills which present verbal propositions to be interpreted⁹ (Lussac & Lussier, 2003, p.53).

Impact of nationality and disadvantaged characteristics on linguistics scores

We have described the scores in this sixth-grade sample by referring to the Reference Table by Flessas and Lussier (2003). In view of the lack of a notable difference between the cohort of exclusively French pupils (90 items of valid data) and the 124 items of valid data that included pupils of foreign nationality, the results for the entire cohort are presented in Table 3. As indicated in the third and fourth column, the standard deviation is rather the same in the two tested populations. It is important to point out though that for the Riddles and the Lexical skills subtests, the maximum score level was not attained.

⁹ for example “*the c..... is a tool used to draw circles*”, where the pupil has to choose the keyword from the list given.

Table 3. Comparison of scores (means and minimum vs. maximum scores) for two cohorts (124 pupils vs. 90 pupils)

Subtest	Mean	Number participants	ofMinimum/Maximum score
Linking words (/13)	5.080 (2.438)	124	0/13
	5.355 (2.395)	90	0/13
Riddles (/12)	6.088 (2.508)	124	1/11
	6.477 (2.473)	90	1/11
Missing words (/11)	6.153 (2.439)	124	0/11
	6.433 (2.248)	90	1/10
Expressions (/5)	2.459 (1.483)	124	0/5
	2.522 (1.508)	90	0/5
Lexical skills (/16)	4.2581 (3.717)	124	0/14
	4.5889 (3.753)	90	0/14

Note. Standard deviations are indicated in brackets

We based our initial results on the sample group of 124 vs. 90 items of valid data, compared with the means and standard deviations corresponding to a sample type of pupils in primary school (based on 55 participants) and middle school (based on 102 participants) in the Reference Tables given by Flessas and Lussier (2003, p. 114¹⁰). As illustrated (Figure 1), there is a very clear adjustment from a low to a high degree of performance. The disadvantaged cohort, with and without foreign nationality, is less competent in all subtests than at a standard middle school. This observation was also noted at the primary level as well.

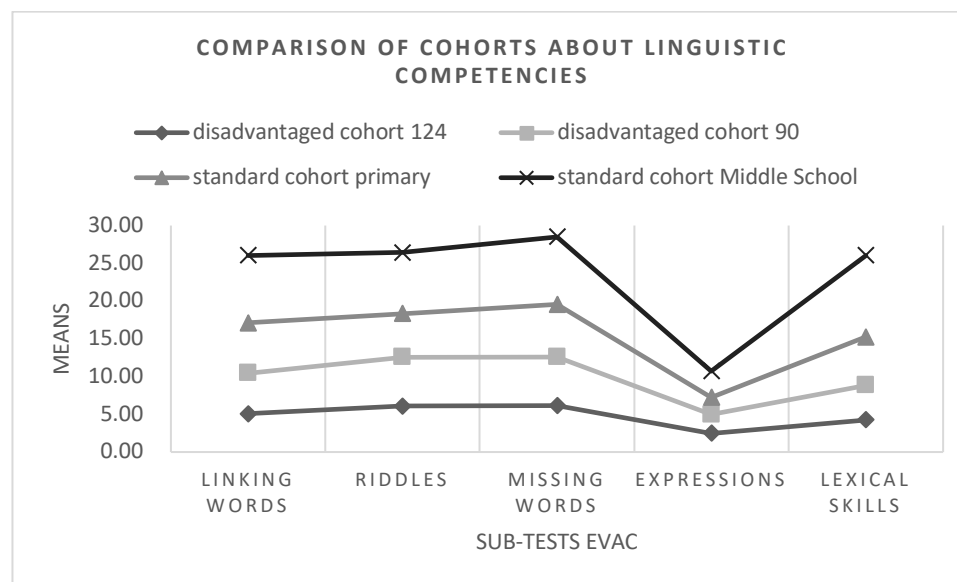
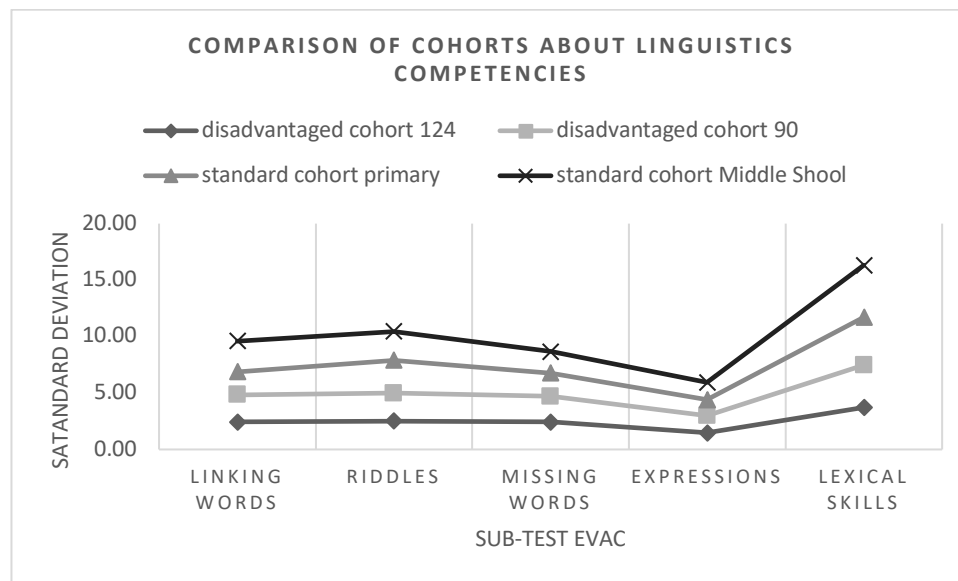


Figure 1. Comparison of means between standard populations (primary and Middle School) and both groups with (124) and without (90) foreign nationality tested in disadvantaged areas

¹⁰ The standard elementary school sample group had the following scores: Riddles: $m=5.78$, $\sigma=2.93$; Missing words: $m=6.98$, $\sigma=2.09$; Expressions: $m=2.25$, $\sigma=1.42$; Linking words: $m=6.69$, $\sigma=2.03$; Lexical skills: $m=6.38$, $\sigma=4.24$. The standard middle school sample group had the following scores: Riddles: $m=8.08$, $\sigma=2.53$; Missing words: $m=8.93$, $\sigma=1.9$; Expressions: $m=3.48$, $\sigma=1.53$; Linking words: $m=8.92$, $\sigma=2.73$; Lexical skills: $m=10.82$, $\sigma=4.61$.

More interestingly, the comparison between standard cohorts and disadvantaged cohorts for both groups (124 vs. 90) indicate that the disadvantaged pupils are less different in their score than the ordinary, standard population (Fig.2). Professors, teaching in disadvantaged areas face disadvantaged pupils and a higher homogeneity of population than in an ordinary teaching situation, which increases the necessity of a heightened adaptation between the lower and higher levels of the pupils (see above, cf. Clanet, 2005).

Figure 2. Comparison of standard deviation about mean scores between standard populations (primary and Middle School) and both groups with (124) and without (90) foreign nationality tested in disadvantaged areas.



Gap between the disadvantaged and standard sample concerning linguistic performance

The subtest of expression seems the sole test without a high degree of differentiation between each cohort (cf. above, Figure 1 and Figure 2), because of a reduced gap in the differentiation of responses (0-5). Considering this fact, it is important to clearly describe the repartition of the population for the graduate percentiles (25th, 50th and 75th) to have another accurate and complementary representation of the disadvantaged sample compared with the standard population. We illustrate this proportion in detail for each subtest: *Riddles*, *Missing Words*, *Linking Words*, *Lexical Skills* and *Expressions*.

Firstly, we illustrate the subtests (Figure 3) by placing the normal curve for the standard level in sixth grade. The global repartition of scores respects the Gaussian curve. It seems the disadvantaged sample corresponds to a retard without abnormal distribution.

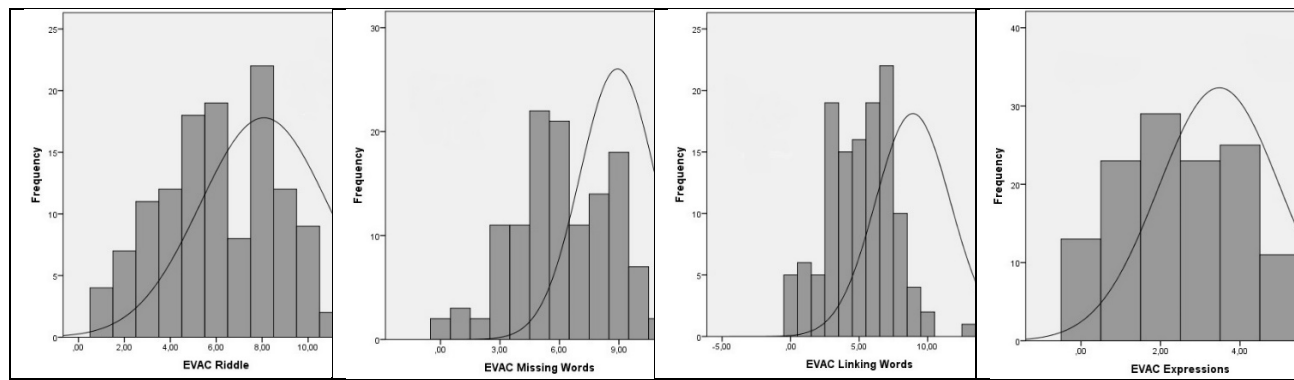


Figure 3. Distributions of scores ($N = 124$) from a) *Riddles* b) the *Missing words* c) the *Linking words* and d) the *Expressions* compared with the standard curves for sixth-grade standardized level ($N=102$, cf Flessas & Lussier, 2003, p. 178)

Riddles subtest

As for Riddles subtest¹¹ 36% of the pupils (score 8 to 11) were above the standard average (50th percentile of the standard reference group) with 8.9% in the top 90%, 9.7% in the top 70% and 17% in the top 55% (55th percentile). On the other hand, 6.5%, 15.3% and 14.5% were in the 45th, 35th and 25th percentiles, respectively. Lastly, the pupils who only solved one out of four riddles were in the bottom 10% of their age group; this was 27.4% of those concerned. Only 42.5% reached the 45th percentile. 1.6% of pupils were in the 90th percentile. None reached the total achievement level represented by 5% of sixth-grade middle school pupils in France.

Missing Words subtest

As for the Missing Words subtest, the pupils had to identify, from four suggestions, the word that fit best into an expression. For this subtest, 21.8% of the pupils were in the top 70% of their age group, while 7.3% were above the 85th percentile. 33.1% were in the average of the top 50%, which placed them at 17% below the standard for the group. For those with lower scores, 22.4% were below the bottom 10th percentile for their reference age, which is 12% more than the average. 17.7% were in the bottom 15th percentile, 16.9% were in the 25th percentile and 8.9% were in the 35th percentile, which meant that 43.5% were in the median range, although some were well below the standard average.

Linking Words subtest

Several *linking words*¹² could be counted as correct. 0.8% of pupils achieved the top score. On the other hand, no pupil was among the top 90% that found 12 or 11 words in standard population). 1.6% were in the 80th percentile and 31.5% were above the 50th percentile. For those with lower scores, 25.1% were below the bottom 10th percentile, which was 15% more than the standard group in this category (0 to 3 correct answers). There were then 15.3%, 12.9%, and 12.1%, who were in the 35th, 20th, and 15th percentiles. In addition, only 46.8% of the pupils were above the bottom 35%, while 60% (compared to the 80% in the reference) were considered as below the bottom 20th percentile of the standard population (five successfully identified the 12 words).

¹¹ The goal of the *riddles* was to identify relatively *common vocabulary words*, such as *cat*, *butterfly*, *rain*, *motorbike*, etc.

¹² The usage of linking words in context corresponds to introducing *where*, *when*, *while*, *how* in sentences: pupils had to find the suitable linking word in sentences where *mmm* was used in place of logical links.

A part of these results is reported to illustrate and compare these three subtests (Fig. 4).

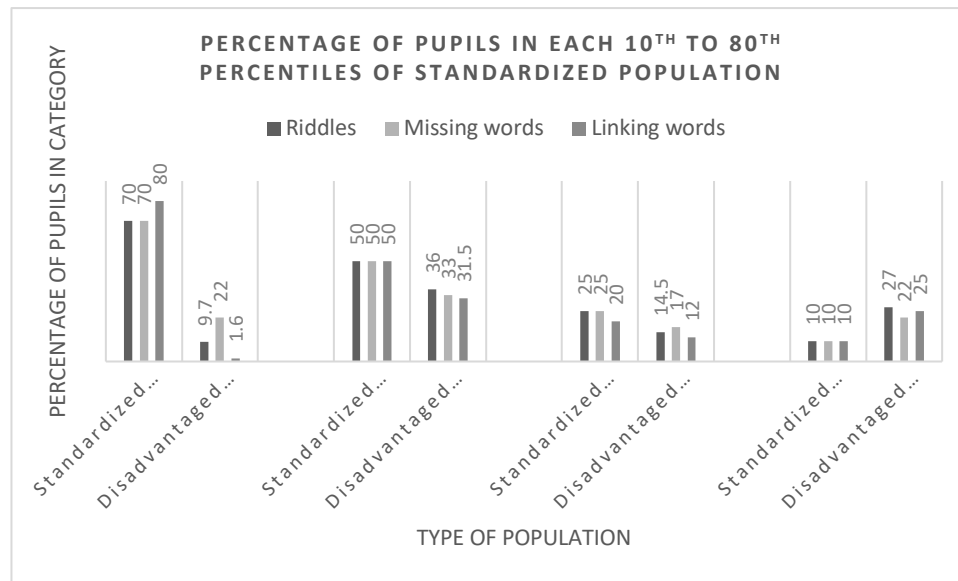


Figure 4. Proportion of pupils ($N=124$) (in % per standardized percentile $N=102$, cf Flessas & Lussier, 2003, p. 178) from a) *Riddles* b) the *Missing words* c) the *Linking words*

As shown in Figure 4, almost a quartile of the disadvantaged pupil sample is cited under the 10th percentile for all of these subtests, around 30% of the disadvantaged pupil sample reaches the median of the 50th percentile of performance reached by the standard population. Dramatically, when 70% or 80% of the standard population performed the three tests, a lower proportion of the disadvantaged pupils represented this possibility: less than 2% in the case of the *Linking Words* subtest. This distribution reflects a phenomenon of deviation from the norm that goes beyond the notion of simple academic delay.

Expressions and Lexical Skills subtests

For the *Expressions* subtest, pupils had to choose the literal meaning of an expression from four options given in a non-literal manner (cultural metaphor). With regard to this subtest, 29% of the pupils were above the 70th percentile, which put them in line with the standard result. 44.6% were in the 50th percentile, putting them in line with the standard result. 18.5% were in the 15th percentile, while 10.5% were in the 10th percentile. For this test, the six-grade pupils from the disadvantaged area sample confirms the idea of a simple academic delay.

The Lexical Skills subtest¹³ presents some confirmation and differences in comparison with other subtests. No pupil was among the top 95% having done well in the test (15 or 16 words found). 0.8% of the sample group were among the top pupils in the 90th percentile with 14 out of 16 words found. 3.2% were in the 80th percentile, instead of the expected 20%, and 3.2% were in the 70th percentile, equaling 8% compared to the expected 30%. Only 28.2% exceeded the 50th percentile, equaling 22% of the pupils below the average for the standard population. 55.6% of the pupils in the sample group were above the 25th percentile; a quarter of the pupils

¹³ Sixteen expressions had to be completed with a word, with the first letter of the word given as a clue, e.g. w for *wrist*, c for *compass*, etc.

of the cohort did not achieve the standard level required. Among the lowest scores, 20.2% and 13.7%, respectively, some were at the expected level of 5%, where 1% of the standard population normally failed at this task.

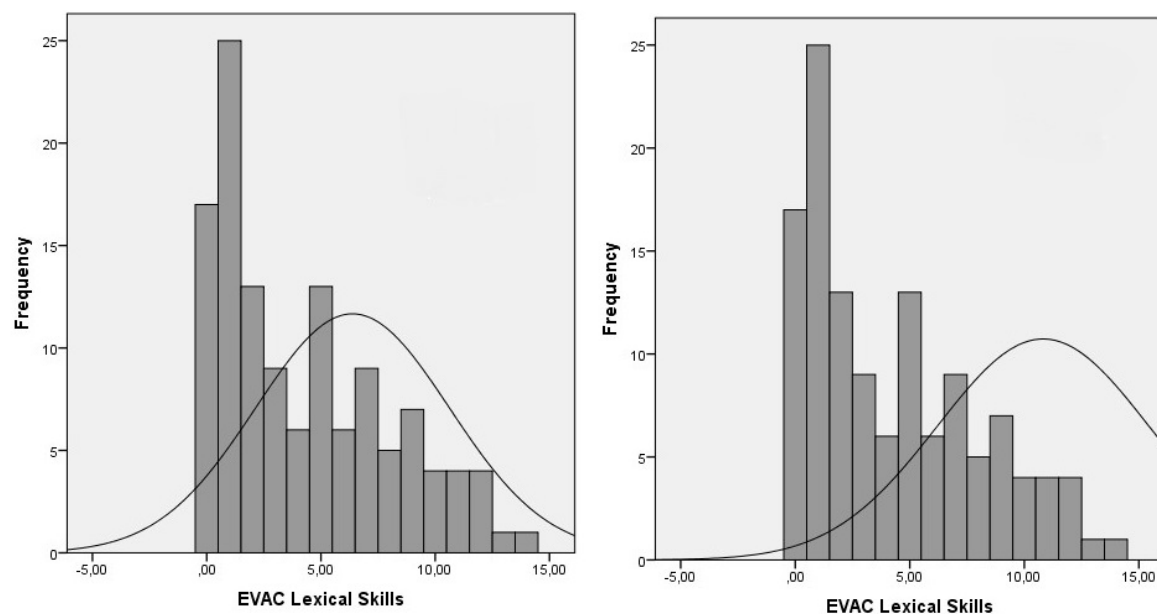


Figure 5. Distribution of scores ($N = 124$) from the *Lexical* skills subtest with the normal Gaussian curve for a) elementary-school level left ($N=55$, Flessas & Lussier: 178) and for b) sixth-grade level right ($N=102$, cf. Flessas & Lussier, p. 178)

This subtest appeared to be the most discriminating, putting the sample group of sixth-grade pupils at schools in disadvantaged areas largely under the average French level.

Peer class effect on fluid intelligence scores

Using the 92 valid data items, cross-tabulated with only the 63 valid data items from the ECL-2 and the EVAC test, the Raven's test tested the homogeneity of the various sixth-grade classes tested in terms of fluid intelligence. For the 63 pupils who took all the tests, general language ability (EVAC) was significantly correlated with that of fluid intelligence ($r=.304$, $p<.003$), whereas, surprisingly, general comprehension (ECL-2) was not significantly correlated with fluid intelligence ($r=.153$, ns). In fact, the easiest series (A) and the penultimate series (D) are not significantly correlated to general scores (EVAC and ECL-2). It is the most difficult series (E) which is doubly correlated with the linguistic aptitude ($r_E=.346$, $p<.001$) and the comprehension ($r_E=.242$, $p<.02$) global scores. In other words, only the best-performing pupils who had good linguistics and comprehension competencies reached a higher degree of fluid intelligence. The result of Raven's Matrices is significantly correlated to the academic level of the pupils in mathematics ($r=.438$, $p<.00001$). Each series of Raven's Matrices (A to E) explain (between 25% to 48%) the contribution of mathematics performance¹⁴. In contrast, the level of French is not in correlation with the general score measuring fluid intelligence. Only the C series of matrices had potential (22%) to be passed (vs. failed) at a significantly higher rate by pupils with a good (vs. poor) level of French ($r_C=.221$, $p<.034$).

¹⁴ The respective correlations are as following: $r_A=.252$, $p<.015$; $r_B=.308$, $p<.003$; $r_C=.484$, $p<.0001$; $r_D=.273$, $p<.008$; $r_E=.360$, $p<.0001$.

Perform to Raven-g Test is significantly correlated ($r=.378$, $p<.0001$) to reach various mean score of classes; means are A=40.78; B=44.61; C=43.82; D=35.20; E=36.74; F=33.27¹⁵, respectively (see Figure 6).

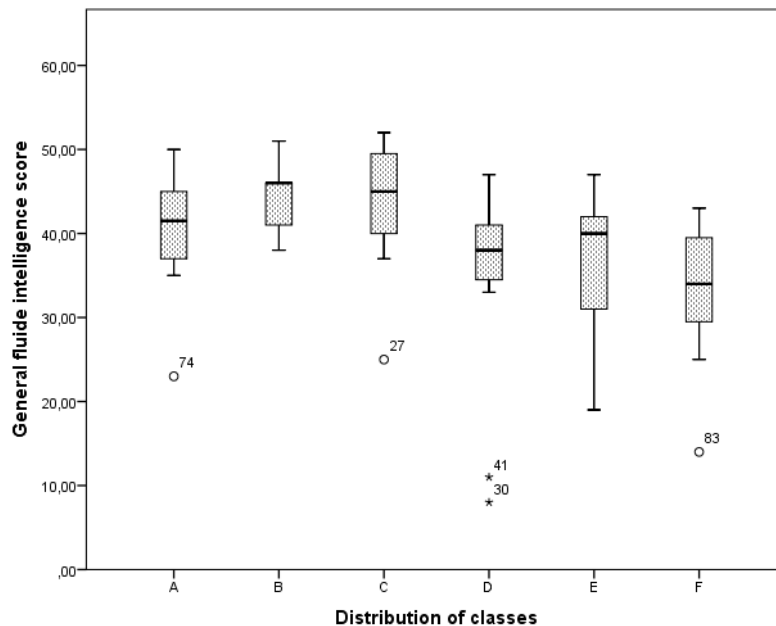


Figure 6. Level of fluid intelligence for each class based on 92 items of valid data

Classes B and C could be distinguished by a very high score placing, surprisingly, the pupils of these classes in the 75th percentile of the standard population of 11-year-olds, so reaching an average score corresponding to at an age of 14.5 years old (cf. Raven et al., 1998, p.61). In contrast, class F remained in the 25th percentile for the same 11-year-olds. Perform to linguistics test (EVAC) and comprehension test (ECL-2), the class C presented a specific profile (Figure 7) with score as respectively following: A=12.83/28.50; B=12.33/27.33; C=7.45/21.91; D=11.33/25.73; E=11.68/24.84 and F=10, 27/20.27. The pupils in class C had one of the lowest averages¹⁶ in the EVAC and ECL-2 tests.

¹⁵ The average general score for 13 year olds was 43 and 44 for 13.5 year olds, reaching 45 for 14 year olds.

¹⁶ The mean for the class is in linguistic aptitude ($m=21.91$) and in comprehension ($m=7.45$).

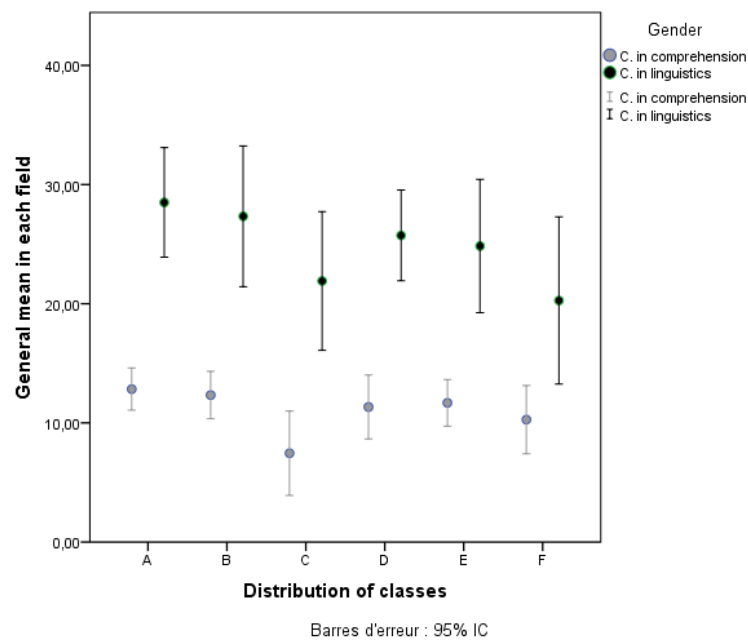


Figure 7. Level of comprehension and linguistics for each class based on 92 items of valid data

As the class C ($m_c=6.6273$) was ranked higher than average in mathematics ($m=6.6228$), a multiple regression analysis test level of mathematics and class, and verified and found a simple trend effect for class ($t(88.4)=1.770$, $p<.08$), no simple effect of level mathematics ($t(88.4)=.806$, ns) and no interaction ($t(88.4)=.994$, ns).

GENERAL DISCUSSION AND CONCLUSION

The study, conducted in a French middle school, had three objectives for a cohort of 124 pupils enrolled in the sixth-grade (including 34 allophones from foreign countries) divided into seven classes: 1. to establish a horizon for training adolescents who are, a priori, academically disadvantaged with regard to various socio-environmental factors, which are, unfortunately, mixed with and blurred by those of the priority education zone, 2. To promote the interest of standardized tests to inform the teachers in charge of these pupils, and 3. To qualify the gap existing between these potentially disadvantaged pupils and all the pupils schooled in other diverse areas (urban, peri-urban, rural) with regard to their academic success and their relationship to knowledge. The ambition of these three objectives was inserted into a particular context: to overcome the reluctance of teachers, in concern for the level of linguistic skills of the adolescents in their care, to implement the practice of philosophical dialogue, despite its recognized benefits.

Our results relate to the scores obtained by pupils in two areas (fluid intelligence and language skills) crossed with academic levels measured by teachers in two subjects (mathematics and French). These measures predict pupil performance (see section I.3.) and are thus used as an indicator of dropout prevention or school failure. The interest of our study lies in the finesse of the measures in order to avoid the sociological bias that reduces the link between the school environment and school failure to a simplistic or sociologically biased causal statement (cf. Dreeben & Barr, 1988). This required comparing scores at three levels: 1. individual specific to each subject, 2. collective associated with the class group (A, B, C, D, E, F and G) and the teacher's

style, 3. the cohort (sixth grade) in terms of standard age -11 years old in France. The conclusion and discussion will focus on the most salient results with regard to these three objectives and this triple plane of scores collected.

Firstly, there is an educational lag of about 20 percentiles characterizing the cohort of adolescent pupils in the sixth grade in a priority education zone compared to the standard middle school population (see Figures 1 and 2). This result confirms the probable sociological bias if this numerical finding alone describes the cohort (cf. Creemers & Reezig, 1996; Sorhagen, 2013). The curves obtained reveal a destabilizing hierarchy of scores: a pupil is all the more likely to obtain a score below the average for his or her cohort if he or she is educated in a disadvantaged area (cf. 90 pupils without foreign nationality) and if he or she comes from an ethnic minority (cf. the 34 pupils with foreign nationality in the cohort). Leaving it at that, we confirm a trend of inequality that has been studied (Bourdieu, 1997; Drebbsen & Barr, 1988), reducing priority education areas to a stereotype of discrimination. This weakens the possibility of teaching differently and adapting an adjusted style of instruction and education (cf. Clot, 2007), with sociological and educational variables remaining confused. In other words, we are signaling the failure of the school in these priority education zones in advance, preventing the redesign of a positive relationship to knowledge with and for these pupils (cf. Charlot, 1999; Charlot et al, 1992; Lahire, 1993; Van Zanten & Fijalkow, 1994).

Secondly, our results refine this observation. The academic lag measured at the level of the cohort of sixth-graders in priority education zones illustrates a gap of about 2 years that could be thought to be catchable: it is indeed a delay, with the standard Gaussian curve serving as a benchmark (cf. Figure 3 illustrating the *Riddles*, *Missing Words*, *Linking Words* and *Expressions* subtests). However, the very distribution of pupils, beyond the mean or median ruling the delay located in the standard percentiles (cf. Figure 4.), reveals notable differences (cf. our §. III.3.1.; III.3.2. and III.3.3.) in the size of the effect (cf. Van Ewijk & Slegers, 2010). The proportion of pupils in the sixth-grade cohort from priority education zones exceeding the 50th percentile is reduced to around 30% for the *Riddles* and *Linking Words* tests, in contrast to the distribution for the *Missing Words* test. This means that, on the one hand, 30% of pupils in priority education zones have an above-average level, which is an undeniable asset for the teacher in conducting adjusted teaching (cf. Clanet, 2005), and on the other hand, that the distribution of pupils in the theoretical categories of good, average and/or weak pupils varies according to the type of language skills test. The teacher, aware of these elements, could probably conduct his teaching differently. In this respect, our results on one of the subtests (*Lexical Skills*, see Figure 5) show a very low percentage of pupils with adequate language skills, with 55% of these adolescents (excluding allophone pupils) producing scores below the 25th percentile (see section III.3.5). The very notion of linguistic competence(s) must be put into perspective and reduced to more targeted sub-competences, in terms of delay size. Teachers would undoubtedly adjust or conduct their teaching better if their pupils were tested beyond academic scores in French. The correlation between academic levels in French, language skills and comprehension skills is effective (cf. Table 1) but, like all correlations, it is circular and not explanatory. In this way, we do not reach the degree of influence that makes the share of each of the factors involved explicit (cf. Osborn, 1997; Tudge & Ducet, 2004), let alone the explanatory share of the interfering factors (cf. Miller & Votruba-Drzal, 2017) responsible for the potential academic delays vs. failures.

Thirdly, our results show that ethnic minority pupils (cf. our 34 foreign pupils) reached the maximum proficiency score on all tests as much as the others (cf. our 90 pupils, Table 3.). In terms of linguistic skills, the gap between pupils of foreign nationality(ies) and French nationality(ies) would thus be less discriminating if it was tested in a standardized manner; the delay of the allophone pupil in mastering the French language should be taken into serious consideration for teaching purposes but should also be relativized in terms of our results.

Fourthly, our results shed light on an important phenomenon: fluid intelligence, considered to be the most robust indicator of academic success (cf. our §. I.3., in particular Bartels et al., 2002; Neisser et al., 1996; Sternberg et al., 2001), would make it possible to discriminate between the scores of pupils at the level of the class group (cf. Figures 6 and 7). Although the conditions of ecological collection did not allow us to delve further into these results (see our §. Limitations), they remain of an important nature, in particular, Matthew Lipman, leader in the development of philosophical dialogue practices intended to raise the academic level of pupils, relied on the level of general intelligence as a developmental factor and lever (see Leleux, 2005). However, it seems that it is precisely on this aspect, measured in a standardized manner in our study (Raven g-scores matrices), that the classes of pupils differ significantly from one another.

Finally, the lack of significant gender differences, which is beyond the scope of our study, is worth noting. Girls tend to outperform boys in the primary school years as well as in high school (Downey & Vogt Yuan, 2005; Fischer et al., 2013), but girls do not outperform boys in general intelligence (Deary et al., 2007). This secondary influencing factor acts on school performance (cf. § II3, in particular Deary et al., 2007) after general intelligence (g-factor, see above). However, as it has been shown that gender differences in achievement are more pronounced among children/adolescents from minority backgrounds (African-Americans vs. Europeans), on the one hand, and that, on the other hand, living in a poor vs. less poor neighborhood also influences school achievement (see Leventhal & Brooks-Gunn, 2004). Our study points out, after all, that the extraction of explanatory and responsible factors for the schooling of pupils in priority education zones requires sociological caution and additional research adapted to national contexts and cultures (in this case, France), linked to the school cartography imposed by the successive ministries of education (see our §. Limitations). Over a three-year follow-up, Leventhal and Brooks-Gunn (2004) show that moving to low-poverty neighborhoods has positive effects on the academic performance of boys aged 11 to 18, compared to their peers living in high-poverty neighborhoods. Sorting out the influencing factors to help teachers better direct their teaching is currently urgent. Priority education 'zones' in France combine a number of influencing factors that our study, after others (e.g., Van Ewijk & Slegers, 2010), helps to reveal. The practice of philosophical dialogue, based on the discriminating character of general intelligence as an essential factor of development (cf. Lipman et al., 1980), and the absence of gender difference associated with this same factor, should perhaps be reconsidered in order to provide, without in any way minimizing the linguistic lag of pupils in the sixth-grade in a priority education zone, and relaying preliminary and complementary studies (cf. Slusarczyk et al., 2015, 2020), more detailed information on how to support middle school teachers.

In conclusion, the specific teaching stage in adolescence could be a key period to help redistribute the cards of school success and fight against inequalities that are as much structural (cf. Bourdieu, 1979, 1989) as functional (cf. the significant differences noted here between school classes): between general intelligence and linguistic competence, the

pedagogical proposal must be refined. At the same time, as long as we do not act in a more regulated and territorially equitable way on the school map, the priority education zones in France, like the disadvantaged geographical, sociological and cultural environments in the USA, will continue to amalgamate the factors that are detrimental to the intellectual success of each nation. Raising the general level of intelligence, which is useful for increasing critical, creative, social and language intelligence (cf. Auriac-Slusarczyk & Maire, 2020; Gorard et al., 2017; House, 2016; Siddiqui et al., 2019), remains the key to a successful school, and the current two-year educational lag of pupils in the sixth-grade grade remains unacceptable.

LIMITATIONS OF THE STUDY

This study, carried out within the framework of an educational project, remained focused on a sector, an educational team, educational choices, and on the results produced, which are not experimental and are not intended to be reproduced. The ecological conditions of the collection must be considered as an operating framework. The implementation of standardized tests provides contextualized results which have no other purpose than to stimulate pedagogical reflection in priority education situations in France. The statistics applied are voluntarily limited to highlighting the effects of the measurements reported and compared at the different levels considered: individual, group and cohort.

References

- Abaspour, N., Ali Nowrozi, R., & Latifi, Z. (2015). Investigating the effect of educating philosophy in the children on the spiritual development of female pupils with 12-14 years old in the city of Isfahan. *Journal of Education and Practice*, 6(11), 162-166.
- Agostini, M. (2012). *Les enfants et la philosophie*. Aux forges de Vulcain.
- Auriac, E (2006). Mesurer l'impact de discussions philosophiques : les tests de raisonnement logique et moraux expérimentés à l'école primaire et en formation d'adulte, *Psychologie et Education*, 2006(2), 39- 57.
- Auriac-Slusarczyk, E., & Maire, H. (2020). Introduction : Sur les pas de Lipman, philosophe à l'école. Une histoire scientifique à connaître. In J.-P. Simon, A. Fournel, S. Lagrange-Lanaspre, & J.-M. Colletta, (Eds.) *Philosopher avec les enfants. Fabrique de l'apprendre. Fabrique du savoir* (pp. 21-59). Presses Universitaires Blaise Pascal.
- Auriac-Slusarczyk, E., Maire, H., Thebault, C., & Slusarczyk, B. (2019). Improving the quality of philosophical writing: Experimentation on the first philosophical compositions of 12 and 13-year old pupils. *Educational Review*, 1-23. doi:10.1080/00131911.2018.1546166
- Auriac-Slusarczyk, E., Thebault, C., & Slusarczyk, B. (2018a)). Mind mapping and philosophy in junior high school. A comparative study of initial philosophical writing in classes of 8th grade pupils used and unused to philosophically directed discussion. *Creative Education*, 9, 1312-1331. <https://doi.org/10.4236/ce.2018.99098>
- Auriac-Slusarczyk, E., Thebault, C., Slusarczyk, B., & Pironom, J. (2018a) Premiers écrits philosophiques à l'école primaire. Productivité conceptuelle et créativité rédactionnelle du CE2 au CM2. *Bulletin de Psychologie*, 71(3), 671-690.
- Auriac-Slusarczyk, E., Adami, J. & Daniel, M.-F. (2011). Tester les prédispositions à l'esprit critique au primaire. *Psychologie et Education* (1), 55-80.
- Ayoub, C., O'Connor, E., Rappolt, G., Vallotton, C., Raikes, H., & Chazan, R. (2009). Cognitive skills performance among young children living in poverty: Risk, change, and the promotive effect of early head start. *Early Childhood Research Quarterly*, 24, 289-305. doi:10.1016/j.jecresq.2009.04.001
- Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Research*, 5, 544-553.

- Bleazby, J. (2020). Fostering moral understanding, moral inquiry & moral habits through philosophy in schools: a Deweyian analysis of Australia's Ethical Understanding curriculum. *Journal of Curriculum Studies*, 52(1), 84-100. <https://doi.org/10.1080/00220272.2019.1650116>
- Brauer, M. (2002). L'analyse des variables indépendantes continues et catégorielles : alternatives à la dichotomisation. *L'Année Psychologique*, 102(3), 449-484.
- Bourdieu, P. (1979). *La distinction. Critique sociale du jugement*. Minuit.
- Bourdieu, (1989). *La noblesse d'Etat. Grande écoles et esprit de corps*. Minuit.
- Cayet, A.-S. (2021). Performer les mots pour philosopher avec des adolescents plurilingues : un tissage de voies/voix. Nouveaux cahiers de la recherche en éducation, Faculté d'éducation de l'Université de Sherbrooke. {hal-03147753}
- Cayet, A.-S. (2020). Le dessin réflexif en philosophie en classe de français langue seconde. In A. Fournel, J.-P., Simon, S. Lagrange-Lanaspre & J.-M. Colletta. (2020). *Philosopher avec les enfants. Fabrique de l'apprendre. Fabrique du savoir* (pp. 77-93). Presses Universitaires Blaise Pascal.
- Cayet, A.-S. (2019). Pour une approche interculturelle sociobiographique en didactique des langues. Expérimentation d'un dispositif réflexif et dialogique : la Discussion à Visée Philosophique (DVP). *Voces y Silencios. Revista Latinoamericana de Educación*, 10 (1), 99-112. {10.18175/vys10.1.2019.08}. {hal-02178440}
- Charlot, B. (1999). *Du rapport au savoir en milieu populaire*. Anthropos.
- Charlot, B., Bautier, E., & Rochex, J.-Y. (1992). *Ecole et savoir dans les banlieues*. Colin.
- Chartier, P., Barbot, H., & Ozenne, R. (2014). L'analyse des erreurs dans les tests de raisonnement logique : principes et illustrations. *L'Orientation Scolaire et Professionnelle*, 43(1). <https://doi.org/10.4000/osp.4296>
- Clanet, J. (2005). Actions/interactions maître-élève et statut scolaire de l'élève. In L. Talbot, *Pratiques d'enseignement et difficultés d'apprentissage* (pp. 35-48). Eres.
- Clot, Y. (2007). De l'analyse des pratiques au développement des métiers. *Education et Didactique*, 1(1), 83-93.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Colom, R., Moriyón, F. G., Magro, C., & Morilla, E. (2014). The long-term impact of Philosophy for Children: a longitudinal study (preliminary results). *Analytic Teaching and Philosophical Praxis*, 35(1), 50-56.
- Creemers, B. P., & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, 7, 197-228. <http://dx.doi.org/10.1080/0924345960070301>
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 13-21. doi:10.1016/j.intell.2006.02.001
- Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *Sociological Quarterly*, 46(2), 299-321. <https://doi.org/10.1111/j.1533-8525.2005.00014.x>
- Dreeben, R., & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, 61, 129-142. <http://dx.doi.org/10.2307/2112622>
- Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 47-69). Russell Sage Foundation.
- Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). Test de Vocabulaire en Images Peabody: TVIP. Circle Pines, MN: American Guidance Service.
- Fouquet-Chauprade, B., Dutrevis, M., & Demeuze, M. (2015). Comparer les PEP : une diversité d'approches pour une diversité de pratiques. *Les politiques d'éducation prioritaire en question, Éducation comparée/ nouvelle série*, 13 7-14. <https://archive-ouverte.unige.ch/unige:88589>.

Fedi, L. (2012). Lipman contre Piaget : une mauvaise querelle à propos de la philosophie pour enfants. *Le Télémaque*, 2(42), 149- 162.

Fiema, G. Slusarczyk, B., Hudry-Critoux, T., Simon, J.-P., Haptel, L., Barroche, I., & Auriac-Slusarczyk, E. (2020). *Analyse des ressentis et perceptions des enseignants quant aux gestes d'improvisation en Communauté de Recherche Philosophique (CRP)*. In A. Fournel, H. Maire, & E. Auriac-Slusarczyk, Gestes et actes d'improvisation dans la praxis de dialogue philosophique, Symposium au Colloque international Improviser l'enseignement, Enseigner l'improvisation, Montpellier, juin 2020.

Fischer, F., Schult, J., & Hell, B. (2013). Sex differences in secondary school success: why female students perform better. *European Journal of Psychology of Education*, 28(2), 529-543.

Flessas, J., & Lussier, F. (2003). *EVAC: Epreuve Verbale d'Aptitudes Cognitives*. ECPA

Fournel, A., Simon, J-P, Lagrange-Lanaspre, S. & Colletta, J-M. (2020). *Philosopher avec les enfants. Fabrique de l'apprendre. Fabrique du savoir*. Presses Universitaires Blaise Pascal.

García-Moriyón, F., Rebollo, I., & Colom, R. (2005). Evaluating philosophy for children: A meta-analysis. *Thinking: The Journal of Philosophy for Children*, 17(4), 14-22.

Giménez-Dasí, M., Quintanilla, L., & Daniel, M.-F. (2013). Improving emotion comprehension and social skills in early childhood through philosophy for children. *Childhood and Philosophy*, 9(17), 63-89.

Gorard, S., Siddiqui, N., & See, B. H. (2017). Can 'Philosophy for Children' improve primary school attainment?. *Journal of Philosophy of Education*, 51(1), 5-22. <https://doi.org/10.1111/1467-9752.12227>

Gregory, M.-R., Haynes, J., & Murris, K. (2016). Philosophy for Children: an educational and philosophical movement. In M. R. Gregory, J. Haynes, & K. Murris, *The Routledge International Handbook of Philosophy for Children* (pp. xxi-xxxi). Routledge.

Harackiewicz, J., Barron, K., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, 94, 562-575. doi:10.1037/0022-0663.94.3.562

Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness Improvement*, 15(2), 177-200.

Heron, G., & Cassidy, C. (2018). Using practical philosophy to enhance the self-regulation of children in secure accommodation. *Emotional and Behavioural Difficulties*, 23(3), 254-269. <https://doi.org/10.1080/13632752.2018.1461469>

Helbling, L.A., Tomasik, M. & Moser, U. (2019). Long-term trajectories of academic performance in the context of social disparities: Longitudinal findings from Switzerland. *Journal of Educational Psychology*, 111(7), 1284-1299.

House, B. (2016). Closing the Achievement Gap with P4C,. *Creative Teaching & Learning*, 6(2), 8-14. https://www.deed.org.uk/wp-content/uploads/2016/02/Creative-Teaching-and-Learning_Feb-2016.pdf

Hughes, J.N., West, S.G., Kim, H., & Bauer, S.S. (2018). Effect of early grade retention on school completion: A prospective study. *Journal of Educational Psychology*, 110(7), 974-991. <https://doi.org/10.1037/edu0000243>

Jenkins, P., & Lyle, S. (2010). Enacting dialogue: The impact of promoting philosophy for children on the literate thinking of identified poor readers, aged 10. *Language and Education*, 24(6), 459-472. <https://doi.org/10.1080/09500782.2010.495781>

Khomsi, A., Nanty, I., Parbeau-Guéno, A. & Pasquet, F. (2005). *ECL-COLLEGE - Évaluation des Compétences Linguistiques écrites au collège*. ECPA. Kuncel, N. R., Hezlett, S. A., & Ones, D. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148-161. doi:10.1037/0022-3514.86.1.148

Lahire, B. (1993). Charlot (Bernard), Bautier (Elisabeth), Rochex (Jean-Yves), Ecole et savoir dans les banlieues. *Revue Française de Sociologie*, 34(4), 690-693.

LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, 41, 55.

Leleux, C. (2014). Discussions à visée philosophique pour développer le jugement normatif des 5 à 13 ans . *Revue Française de Pédagogie*, 186, 75-84.

Leleux, C. (2005). *La philosophie pour enfants. Le modèle de Matthew Lipman en discussion*. De Boeck & Larcier.

Leventhal, T., & Brooks-Gunn, J. (2004). A randomized study of neighborhood effects on low-income children's educational outcomes. *Developmental Psychology*, 40(4), 488-507.

Lévine, J. (2005). Essai sur le monde philosophique de l'enfant. Le dialogue Moi Monde-Instance Monde. In C. Leleux, *La philosophie pour enfants: Le modèle de Matthew Lipman en discussion* (pp.71-93). De Boeck.

Lipman, M. (2003). *Thinking in education*. Cambridge University Press.

Lipman, M., Sharp, A. M., & Oscanyan, F. S. (1980). *Philosophy in the classroom*. Temple University Press.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, 26, 33-51. <http://dx.doi.org/10.1016/j.econedurev.2005.09.008>

Maire, H., Auriac-Slusarczyk, E. & Pironom, J. (2019). Faire écrire des élèves dans une visée philosophique : scénario pédagogique et regard du psychologue. *Studia Universitatis Babeş-Bolyai*, 64(3), 91-117. <https://doi.org/10.24193/subbphil.2019.3.05>

Maire, H., Auriac-Slusarczyk, E., Slusarczyk, B., Daniel, M-F., & Thebault, C. (2018), Does one stand to gain by combining Art with Philosophy? A study of fourth-year college (13/14 years of age)? Philosophical writings produced within the PreCPhi/Philosophemes Corpus. *Journal of Education and Learning*, 7(4). <https://doi.org/10.5539/jel.v7n4p1>

Miller, P. & Votruba-Drzal, E. (2017), Pre-K Classroom-Economic Composition and Children's Early Academic Development, *Journal of Educational Psychology*, 109(2), 149-165.

Millett, S., & Tapper, A. (2012). Benefits of collaborative philosophical inquiry in schools. *Educational Philosophy and Theory*, 44(5), 546-567. <https://doi.org/10.1111/j.1469-5812.2010.00727.x>

Mortier, E. (2005). Études d'évaluation : la méthode de Matthew Lipman comme moyen de développement. In C. Leleux (dir.), *La philosophie pour enfants. Le modèle de Matthew Lipman en discussion* (pp. 47-69). De Boeck.

Morton B. Brown & Alan B. Forsythe, (1974). Robust Tests for the Equality of Variances, *Journal of the American Statistical Association*, 69(346), 1974, 364-367

Murphy, K.-P., Wilkinson, I. A.-G., Soter, O.-A., Hennessey, N.-M., & Alexander, F.-J. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, 101(3), 740-764. <https://doi.org/10.1037/a0015576> Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.

Neuman, Y. (2003). Go ahead, prove that God does not exist! On high school student's ability to deal with fallacious arguments. *Learning and Instruction*, 13(4), 367-380.

Osborne J.W. (1997). Race and academic disidentification. *Journal of Educational Psychology*, 89(4), 728-735.

Plumelle, B. (2016). Quelle réussite scolaire pour les élèves issus de l'immigration ? Politiques, dispositifs, préconisations. Centre international de recherches documentaires. <https://www.ciep.fr/sites/default/files/atoms/files/focus-quelle-reussite-scolaire-pour-les-eleves-issus-de-l-immigration-politiques-dispositifs-preconisations.pdf>

Raven, J. (1998). *PM - Matrices Progressives de Raven. Test de référence pour l'évaluation du facteur g*. ECPA/Pearson.

Raven, J., Court, J.-H., & Raven, J.-C. (1998). *Manuel des Raven. Section 3. Progressives matrices standards (PM38)*, traduit de l'anglais par L. Marcou. Pearson : ECPA.

Rey, V., Poinso, F., De Martino, S., Sabatier, C., & Girardot, A. M. (2007). Acquisition du traitement de la métaphore chez l'enfant de 7 à 10 ans. Quelques propositions. *Glossa*, 99, 4-11.

Säre, E., Luik, P., & Tulviste, T. (2016). Improving pre-schoolers' reasoning skills using the philosophy for children programme. *Trames : Journal of the Humanities and Social Sciences*, 20(3), 273-295.

- Sasseville, M. (1994). Self-esteem, logical skills and philosophy for children, *Thinking*, 4(2), 30-32.
- Siddiqui, N., Gorard, S., & See, B. H. (2019). Can programs like Philosophy for Children help schools to look beyond academic attainment?. *Educational Review*, 71(2), 146-165.
<https://doi.org/10.1080/00131911.2017.1400948>
- Slusarczyk, B. (2020). L'intelligence collective lors du délicat passage de l'adolescence au collège, In J.-P. Simon, A. Fournel, S. Lagrange-Lanaspre, & J.-M. Colletta (Eds.), *Philosopher avec les enfants. Fabrique de l'apprendre. Fabrique du savoir*. (pp. 61-76). Presses Universitaires Blaise Pascal.
- Slusarczyk, B., Fiema, G., & Auriel, A. & Auriac-Slusarczyk, E. (2015). Etude de l'impact d'une introduction de la philosophie dans les curriculums au primaire et au collège sur l'intégrité cognitive. *Recherche et Education*, 14, 123-145.
- Slusarczyk, S., Rey, V., Romain, C. & Auriac-Slusarczyk, E. (2020). La compréhension des métaphores de 8 à 14 ans. In J.-P. Simon, A. Fournel, S. Lagrange-Lanaspre & J.-M. Colletta, (Eds.). *Philosopher avec les enfants. Fabrique de l'apprendre. Fabrique du savoir* (pp. 121-158). Presses Universitaires Blaise Pascal.
- Sorhagen, N.S. (2013). Early teacher expectations disproportionately affect poor children's high school performance, *Journal of Educational Psychology*, 105(2), 465-477.
- Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, 47(1), 1-41.
- Topping, K. J., & Trickey, S. (2007). Collaborative philosophical inquiry for school children: cognitive gains at 2-year follow-up. *British Journal of Educational Psychology*, 77(4), 787-796.
<https://doi.org/10.1348/000709907X193032>
- Topping, K. J., & Trickey, S. (2014). The role of dialog in philosophy for children. *International Journal of Educational Research*, 63, 69-78. <https://doi.org/10.1016/j.ijer.2013.01.002>
- Trickey, S., & Topping, K. J. (2004). Philosophy for children: a systematic review. *Research Papers in Education*, 19(3), 365-380. <https://doi.org/10.1080/0267152042000248016>
- Trickey, S., & Topping, K. J. (2006). Collaborative philosophical enquiry for school children: socio-emotional effects at 11 to 12 years. *School Psychology International*, 27(5), 599-614.
<https://doi.org/10.1177/0143034306073417>
- Tudge, J. R., & Doucet, F. (2004). Early mathematical experiences: Observing young Black and White children's everyday activities. *Early Childhood Research Quarterly*, 19, 21-39.
<http://dx.doi.org/10.1016/j.jecresq.2004.01.007>
- Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5, 134- 150. <http://dx.doi.org/10.1016/j.edurev.2010.02.001>
- Van Zanten, A., & Fijalkow, J. (1994). Débat autour d'un livre : Charlot (Bernard), Bautier (Elisabeth), Rochex (Jean-Yves), École et savoir dans les banlieues... et ailleurs. *Revue Française de Pédagogie*, 106, 111-116.
- Walker, C. M., Wartenberg, T. E., & Winner, E. (2013). Engagement in philosophical dialogue facilitates children's reasoning about subjectivity. *Developmental Psychology*, 49(7), 1338-1347. <https://doi.org/10.1037/a0029870>
- Yan, S., Walters, L. M., Wang, Z., & Wang, C. C. (2018). Meta-analysis of the effectiveness of philosophy for children programs on students' cognitive outcomes. *Analytic Teaching and Philosophical Praxis*, 39(1), 13-33.