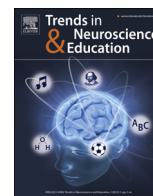




ELSEVIER

Contents lists available at ScienceDirect

Trends in Neuroscience and Education

journal homepage: www.elsevier.com/locate/tine

Research Article

How formal education and literacy impact on the content and structure of semantic categories



Régine Kolinsky^{a,b,*}, Rosemeire Selma Monteiro-Plantin^c, Elias José Mengarda^d, Loni Grimm-Cabral^e, Leonor Scliar-Cabral^{e,f}, José Morais^b

^a Fonds de la Recherche Scientifique–FNRS, Belgium

^b Unité de Recherche en Neurosciences Cognitives (UNESCOG), Center for Research in Cognition & Neurosciences (CRCN), Université Libre de Bruxelles (ULB), CP 191, 50 Ave. F. Roosevelt, B-1050 Brussels Belgium

^c Universidade Federal do Ceará, Brazil

^d Universidade Federal de Santa Maria, RS, Brazil

^e Universidade Federal de Santa Catarina, UFSC, Brazil

^f Conselho Nacional de Pesquisa, CNPq, Brazil

ARTICLE INFO

Article history:

Received 12 November 2013

Received in revised form

31 May 2014

Accepted 22 August 2014

Keywords:

Literacy effects

Formal education effects

Semantic categorization

Semantic fluency

ABSTRACT

We examined the hypothesis that formal education and literacy impact the richness and precision of semantic knowledge but not the organization of semantic categories and basic mechanisms of access to them.

In Experiment 1, adults of varying levels of formal education were presented with semantic fluency tests and a superordinate naming task. Experiment 2 examined the impact of reading proficiency on adults of varying degrees of literacy. They were presented with simple semantic, alternating semantic and phonemic fluency tasks, as well as with literacy-related, reasoning and memory tests.

Fluency was analyzed in terms of overall performance, sequential order and speed of responses. Despite lower performance, illiterates and adults with null or limited formal education displayed taxonomic clustering and retrieval by semantic subcategory, as did participants with higher formal education levels. Yet, formal education and literacy slightly speed up access to categories, probably providing useful cues for generating category exemplars.

© 2014 Elsevier GmbH. All rights reserved.

1. Introduction

A key to concept acquisition is the ability to categorize objects, namely, to establish that a perceived entity belongs to a particular group of things that share some important characteristics. Both taxonomic and thematic relations between objects are considered as essential in this process, as both allow drawing inferences, namely, deriving information not explicitly present in an experiential object. By treating non-identical entities as if they were the same (e.g., both terriers and Golden Retrievers are dogs), taxonomic relations support inferential generalizations from one entity to another, non-identical, entity (e.g., if terriers are dogs and have a heart, Golden Retrievers, as they are also dogs, also have a heart). Furthermore, thanks to their hierarchical structure and to the

principle of class inclusion, taxonomic categories allow facts known about higher-level concepts to apply to lower ones (e.g., learning that animals have a heart and learning that alligators are animals allows inferring that alligators have a heart). On their side, by providing “conceptual coherence” [28], thematic relations occurring between entities independently of their intrinsic nature (e.g., functional and spatial relations) help to generate expectations about events or scenarios (see also [14,55]). For instance, although a leash is not a kind of animal, the fact that leashes are used on animals may help identifying something on a leash as being an animal [28]. According to recent evidence (e.g., [29]), functional features might even be privileged in biological kind classification.

Much recent work focused on how expertise and culture modulate the relative use of thematic vs. taxonomic relations (e.g., [3,37]). According to Estes et al.’s review [14], although formal education appears to discourage thematic thinking (e.g., [30,60,61]), this relationship may vary across cultures. Western cultures emphasize taxonomies by attending to objects and attributes, whereas East Asian cultures emphasize themes by attending to relations and contexts (e.g., [43]).

* Corresponding author at: Unité de Recherche en Neurosciences Cognitives (UNESCOG), Center for Research in Cognition & Neurosciences (CRCN), Université Libre de Bruxelles (ULB), CP 191, 50 Ave. F. Roosevelt, B-1050 Brussels Belgium. Tel.: +32 26504230.

E-mail address: rkolins@ulb.ac.be (R. Kolinsky).

Less attention has been paid to the influence of formal education, including literacy, on the structure of semantic categories and on the mechanisms of access to them. This is probably due to the fact that young, preliterate and/or preschool, children already show a rich organization of semantic knowledge. Indeed, taxonomic capacities appear very early in development. Three-year olds understand the logic of class inclusion, for instance that *car* and *bike* belong to the superordinate *vehicle* [33]. Four-year olds can use both thematic (e.g., *zoo–elephant*) and superordinate (e.g., *animal–elephant*) relations as recall cues very efficiently [6], and preliterate six-year olds display an auditory semantic priming effect when primes and targets are either coordinates (e.g., *arm–leg*) or hierarchically related (e.g., *fruit–apple*) [46]. Furthermore, in five-year olds, categorical relations lead to more false recognitions of a probe in a sentence than the part–whole relationship [32]. Notably, these children are more affected when the probe is a superordinate than when it is a subordinate. This directional effect argues against the interpretation that they simply learned category names as synonyms (e.g., *Lassie, dog* and *animal* as alternative names).

Nevertheless, none of these findings disproves a significant impact of formal education and literacy on semantic knowledge. In the present study, we examined the hypothesis according to which semantic representations and processing strategies are qualitatively the same whatever the formal education and literacy level, despite the fact that the richness and precision of knowledge (henceforth, *content knowledge*) are augmented by increased education and literacy. Under this view, learning in the classroom, and in particular science learning, as well as activities linked to literacy (reading books, magazines, etc.), would thus only impact on content knowledge but not on the structure of semantic categories, including their hierarchical organization, or on the basic mechanisms of access to these categories.

Though seemingly simple, this idea is difficult to examine. In developmental studies, literate children with higher levels of formal education are typically older than children who are preliterate and have received less formal education, leading to confound with neural maturation or other cognitive abilities developing with age (e.g., executive functions, including cognitive flexibility [22]). Actually, estimating the effect of formal education and literacy requires comparing groups of adults of varying levels of formal education and literacy, including, ideally, adults who are completely illiterate as they have never attended school due to socioeconomic reasons.

Several studies examined such populations in semantic fluency tasks, in which participants are asked to generate words belonging to a specific taxonomic category, for instance animals. This task is widely used in cognitive psychology and neuropsychological assessment because it provides useful information about both retrieval strategies and lexico-semantic networks [42,63]. Indeed, although words are not concepts in any simple sense (e.g., [31]), sequences of word retrieval (e.g., listing animal names) are taken as reflecting the way entities are represented in conceptual memory, including their hierarchical taxonomic organization.

Among the various measures of fluency performance, the number of words generated obviously depends on content knowledge, and several studies showed indeed that people having no formal education (in all studies, illiterate unschooled adults) provide far less responses than formally educated people in semantic fluency tasks [8,10,17,23,48,50,62] (see [34] for similar results on unschooled vs. age-matched schooled children). Although both the ecological or cultural relevance of the chosen semantic criteria and the level of reference to concrete knowledge and specific situations modulate performance [17,51,57,56], an impact of formal education has been reported even in an ecologically relevant, concrete version of the task [62], when participants were

asked to name things or food items that can be bought at the supermarket. Other measures that depend on content knowledge also vary with level of education. This is the case for instance of the number of subcategories generated for animals, e.g., pets, insects, birds, fishes, reptiles, etc. [8,17]. In addition to these quantitative differences, some qualitative differences may be expected, as different levels of education are associated with different social-cultural backgrounds. This may affect typicality responses, reflected in fluency tasks by the ordering of members within categories, with more typical members named first [21,53]. In children, it has been shown that socioeconomic status and associated specific experiences influence typicality responses, at least with categories that are related to exposure to formal contexts, such as animals [74]. Similarly, illiterate adults do not mention the same prototypical animal as participants with higher levels of education [8].

Whether formal education and literacy also impact on the very structure and mechanisms of access to semantic categories has seldom been explored and seems to depend on the aspects of performance examined in fluency tasks. When participants are asked to freely generate a list of words corresponding to a given taxonomic category, e.g., animals, they tend to produce these words in clusters, namely in groups of words belonging to the same subcategory, like pets, insects, birds, water animals, etc. [7,18]. This *clustering* component involved in fluency reflects the search in memory for relevant subcategories. It is dissociable from another aspect of performance, called *switching*, which has been defined as the ability to search intentionally for new subcategories and switch effectively from one subcategory to another once a subcategory is exhausted [69]. This requires specific cognitive control processes indicative of flexibility and identified as *shifting* [39], and hence is linked to executive processes [69,70]. For formally educated individuals, switching and clustering would contribute equally to performance in semantic fluency, and would be independent of each other [67,69].

Up to now, the data on the influence of formal education and literacy on these two components of fluency performance are scarce and somewhat abstruse. Clustering may be estimated through several measures. Among them, cluster size (i.e., the number of consecutively generated words belonging to the same semantic sub-category) would measure the ability to produce related words within a given semantic subcategory [69]. This measure has been reported to be similar in illiterate and literate (schooled) adults [17,23], despite the fact that overall performance was augmented by increased education (see also [24]). Results are less clear as regards the number of clusters, which was observed to vary with formal education level in some studies [17] but not in others [8], and the number of isolated words, which in one study [17] was reported as being higher in literates compared to illiterates. As regards the number of switches between subcategories, it seems to be lower in illiterates than in literates [8,17,24], although the effect of formal education on this measure may vary as a function of category, being for instance stronger for animals than for fruits [56]. However, the impact of formal education and literacy on executive processes, in particular on shifting, remains unclear, as it has not always been observed [40].

In addition, there is a debate on which measures are more appropriate to examine the clustering and switching components. Two main types of methodological problems have been identified. The first one is that some of the measures strongly depend on total output. This is the case for instance of switching [13], traditionally calculated as the number of transitions between clusters. To avoid this problem, Tröster et al. [66] proposed to calculate the proportion of switches, i.e., the number of switches adjusted on total output. Yet, the significance and usefulness of this measure has been questioned (e.g., [66,68]), including because it can lead to floor effects and distortion of the data for low output [1]. Although

this has rarely been debated, a similar problem arises with all other measures correlated with total output. In fact, for such measures, it would be more appropriate to statistically control for total output by computing partial correlations (or run covariance analyses) [1].

Second, as a consequence of the timed nature of the fluency task, switching and clustering are mutually exclusive [1] and hence their measures are negatively correlated with each other (e.g., [69]). Thus, as commented on by Abwender et al. [1, p. 326–327], switching might actually “reflect nothing more than the absence of the ability to cluster (i.e., extremely restricted access to semantic subcategories), especially if the switch occurs between two nonclustered words”. This is a problem in several studies, including with illiterates adults [8,17], in which the number of switches was calculated based on the total number of transitions between clusters, including single words. According to Abwender et al., only *cluster switches* (CS), namely the transitions between adjacent clusters defined in their study as including at least two words (and not *hard switches*, i.e., the transitions between clusters and single-word responses or between two single-word responses) would reflect executive processes, in particular mental flexibility.

In addition to these methodological problems linked to the scoring procedure, other researchers have a different view on how semantic processes and executive control affect semantic fluency. For instance, Mayr and Kliegl [36] suggested that the time for every act of retrieval contains two additive components: one executive and the other semantic in nature. The executive component reflects aspects like updating the current search criterion and stopping or initiating single retrieval processes. It is assumed to be relatively constant, no matter whether it occurs within vs. between clusters, or early vs. late in the recall sequence. In contrast, the semantic component reflects the actual semantic search demands, including relatively automatic spreading activation in the semantic network. By this view, between-cluster retrieval should simply require more of the same semantic processing than within-cluster retrieval.

Irrespective of which model will ultimately prove correct, a better way to examine semantic processing would thus be to consider also time measures (see also [35]). For instance, the number of switches not only depends on the difficulty with accessing a new semantic cluster, but is also reduced in case of difficulties to come up with new words within clusters: the more time spent for retrieval within a cluster, the less time remains for accessing new clusters. Hence, a group difference in number of switches could imply that what is prolonged is either the duration of each switch or the duration of each retrieval (no matter whether it occurs within or between clusters). It would thus be interesting to compare the retrieval times for the words generated within a cluster with the retrieval times between clusters. Critically, in formally educated people, the former are shorter than the latter, reflecting retrieval by subcategory (e.g., [18,22]; see reviews in [36,75]). Yet, a more pronounced switching time cost might reflect specific difficulties linked to executive processes.

In addition, computing the retrieval-position function, namely the interword retrieval time as a function of the position of a word within the recall sequence (independently of the fact that it is an intra- or inter-category response), may provide a relatively pure estimate of semantic difficulty. In particular, the slope of the search function may be an indicator of semantic processing efficiency [36], as there is a very well-known reduction in the production rate as a function of time [7] (for a review, see [75]), which is observed in all education groups, even in illiterate adults [8]. According to Mayr and Kliegl [36], demands on semantic processing proper would increase as a function of recall position, whereas demands on executive processes should stay relatively constant across recall positions.

Furthermore, it may be useful to also distinguish *task-consistent* and *task-discrepant* clustering. Indeed, many studies have shown that two types of clusters, phonemic and semantic, are found in semantic fluency (e.g., [26,47]). Abwender et al. [1] argued that task-discrepant clustering (phonemic clustering on semantic fluency), even if less frequent, reflects a more deliberate strategy than task-consistent clustering (semantic clustering on semantic fluency), which could reflect a more passive access to words through automatic activation of nearby concepts (i.e., by spreading activation). Examining both clustering types would be particularly interesting as regards the potential impact of literacy, which is known to be tightly linked to the development of the explicit representations of speech [41] that probably subtend phonemic clustering. This is suggested by the fact that the effect of education and associated literacy level on overall performance is much stronger in phonemic compared to semantic fluency [23,48], namely when participants have to produce words beginning with a given letter or phoneme (e.g., [1,64]).

Finally, other specific components may be important for retrieval in addition to activation spreading from the cue to related items and self-generation of category cues (e.g., phonemic cues) to access new items. In particular, working memory would be important not only to access new items but also to monitor the products of retrieval (i.e., generated items) and suppress previously recalled items, which is essential to prevent errors, especially repetitions [54,71]. Examining the monitoring component through the estimation of repetition errors has rarely been done in studies examining the effect of formal education. Yet, one study observed no significant difference in number of repetitions between illiterate and literate adults [17].

In the present study, we examined these various components of semantic fluency performance in two experiments that included participants with no or limited formal education. Both experiments were run in Brazil, in which low levels of formal education still characterize an important part of the population.¹ Hence, importantly, it was only for social reasons that these participants either had never attended school or had attended school for only a few years. None of them had a history of special difficulty other than lack of access to schools or need to work early in life, and all, including illiterates, were fully socially functional and, except for some housewives, had regular employment.

In Experiment 1, we contrasted two groups of participants that differed by their level of formal education. In one of the groups, participants had no or limited formal education, namely they were either entirely unschooled (and illiterate) or minimally schooled, with up to four years of formal education (henceforth, *-FE* participants); in the other, participants had between 11 and 14 years of formal education, hence all had completed the second level of secondary school (henceforth, *+FE* participants). The two groups were presented with five tests of semantic fluency (animals, working tools, furniture, domestic utensils, and vehicles) as well as with a *superordinate naming* task that may be considered as a mirrored situation of the semantic fluency task. Indeed, in the latter, participants had to provide a superordinate name (e.g., to answer *furniture*) for three orally presented coordinate words (e.g., *sofa-table-cupboard*). This allowed us to further test the hypothesis that people with no or limited formal education lack detailed and precise taxonomic knowledge and/or associated verbal labels.

Experiment 2 was aimed at addressing an additional methodological difficulty, which is to isolate the specific contribution of reading proficiency to semantic knowledge and/or processes.

¹ For example, Yassuda et al. [76] mention a national survey reported in a Brazilian publication [58] according to which 57% of old adults in Brazil have only four years of education and 18% did never attend school.

Indeed, all the studies having examined the impact of formal education and literacy in this field have contrasted illiterate, unschooled people to formally educated, literate people, thereby unconfounding formal education and literacy level. Yet, literacy per se may impact semantic knowledge; for instance, it has been suggested that, in children, greater semantic knowledge leads to better reading comprehension, and vice versa [4]. In Experiment 2, we examined the relation between reading proficiency and the various facets of semantic fluency on a fresh group of participants who had never or almost never attended school in childhood but presented varying degrees of rudimentary reading decoding, as most of them were attending the first class in a literacy program for adults. Through various literacy-related measures, these participants were carefully evaluated as regards their literacy level, and were presented with two semantic fluency tasks. The first was an animal fluency task similar to the one of Experiment 1. The second was an *alternating fluency* task, in which participants had to produce in alternation the name of a fruit and the name of a tool. The latter task allowed further investigating the impact of literacy on the switching component of fluency performance. In addition, the participants of Experiment 2 were presented with a phonemic fluency task, in which performance is supposed to be strongly linked to literacy [23,48]. They were also presented with several tasks aimed at assessing reasoning and memory, especially working memory, which has been argued to be important for verbal fluency performance (e.g., [54,71]).

2. Experiment 1: comparing adults of various levels of formal education

In this experiment, we presented –FE and +FE participants with semantic fluency tests as well as with a superordinate naming task. In five semantic fluency tests, participants were asked to produce names first of animals, next of working tools, next of furniture, next of domestic utensils, and finally of vehicles. For the animal category (which elicited the greatest number of responses and demonstrated to be the category with the least membership ambiguity), in addition to the total number of words generated, we also examined the clustering and switching components.

According to our hypothesis, we expected –FE participants to differ from +FE ones for all measures that depend on content knowledge. In the fluency task, we thus expected –FE to present a lower number of words produced per category and a lower number of subcategories compared to +FE. We also expected the two groups to present different typicality patterns. In the superordinate naming task, we expected –FE participants to produce a lower proportion of *near superordinate* responses, namely of responses as *birds* to the triplet *canary–parrot–swallow*, which are less generic than *distant superordinate* responses such as *animals* to the same triplet. On the contrary, we expected –FE not to differ strongly from +FE participants as regards both the mechanisms of access to taxonomic knowledge and the hierarchical organization of semantic categories. In particular, we expected them to also display a hierarchical organization of semantic categories and hence semantic clusters of similar size as +FE participants, and to present shorter interword time intervals (henceforth, *interword times*) for consecutive responses belonging to the same subcategory (e.g., between two exemplars of birds) than for consecutive responses belonging to different subcategories (e.g., between an exemplar of a bird and an exemplar of a fish). Although this would reflect retrieval by subcategory in both groups, we expected these to differ as regards other, more strategic cues that can be used to access new exemplars, in particular grapho-phonological cues that allow phonemic

clustering. As regards switching, predictions were less clear, as the impact of formal education and literacy on executive processes, in particular on shifting, remains uncertain and has not always been observed [40].

2.1. Method

2.1.1. Participants

The whole sample included 31 adults living in three different cities of the Southern part of Brazil (Palhoça and Florianópolis, in the State of Santa Catarina, and Maringá, in the State of Paraná) who were assigned to two groups, according to their level of formal education. All were Brazilian natives except for one who had arrived in Brazil at two years of age; all spoke only Brazilian Portuguese.

Among the 13 –FE participants (7 women from 25 to 59 years; mean age 40.3 years), seven were totally illiterate and almost unschooled. None of these *illiterates* could name a single letter. Three had never attended school in childhood, and four had attended school only for some months and in an irregular way. The six other participants had attended elementary school in childhood for one up to four years (on average, 3 years). We will refer to them as *semi-literates*, as they only learned to read up to a very rudimentary level, and their decoding ability was slow and sometimes inaccurate. As a matter of fact, presented with a list of 22 pseudo-words, none of them could read accurately more than 15 (68%). These –FE participants had no or limited formal education for social reasons, with no history of special difficulty other than lack of access to schools or need to work early in life. All of them, including the fully illiterates, were fully socially functional and, except for four housewives (2 illiterates, 2 semi-literates), were working as workmen, servants, hairdresser, cook, night guard, or concierge. At the time of testing, nine participants of this group (three illiterates, all semi-literates) were beginning to attend special literacy classes for adults.

The 18 +FE participants (4 women; from 24 to 60 years; mean age: 38.5 years; no significant age difference with –FE participants: Mann–Whitney $U=106.5$, $p > 0.10$) had between 11 years and 14 years of formal education, hence had at least completed the second level of secondary school; nine had completed the third (and last) grade of undergraduate course and were attending a college at the time of testing. Except for one retired participant and one housewife, all were working in administrative jobs, or as teachers, drivers or nurses.

2.1.2. Materials, procedure and scoring

2.1.2.1. Semantic fluency. The experimenter, using a chronometer, asked participants to produce for two minutes as many words as possible, without repetition, corresponding to instances of a specific taxonomic category that was presented orally through its superordinate name. Participants were first asked to produce names of animals, next of working tools, next of furniture, next of domestic utensils, and finally of vehicles. The participant was stopped after two minutes. The whole session was recorded for further data analysis.

Each participant's performance was classified and analyzed as follows:

Total correct output: the total number of correct words generated per taxonomic category (animals, working tools, furniture, domestic utensils, vehicles), excluding errors and repetitions.

Typical members: separately for each category, we attributed the scores of 3, 2 and 1 to the first, second and third items indicated by each participant, and disregarded the items mentioned after the first three. Then, we calculated for each group and category the typicality score of the three most cited items by dividing the total score by the maximum score possible (39 vs. 54

for a given item, for the 13 and 18 –FE and +FE participants, respectively, if all the participants mentioned this item in the first place).

The other measures were only calculated for the animal fluency test.

Number of subcategories: following the suggestion of other researchers (e.g., [52,69]), the semantic subcategories produced were empirically defined based on the words generated by the participants, as pets or farm mammals, wild mammals, birds, reptiles and batrachians, fishes and crustaceans, and insects, arachnids and worms.

Number of semantic clusters: to be sure that a clustering strategy was used, as suggested by Robert et al. [52] and as done in former work on participants with low levels of formal education [8,17,23,56], we only considered clusters of three or more consecutive words belonging to the same semantic subcategory. In doing this, we excluded names of a subcategory for which the participant had already generated a word (e.g., *bird* after *canary*) and intra-cluster repetitions (e.g., the sequence *jaguar-monkey-Jaguar* was not considered as a cluster). However, inter-cluster repetitions were taken into account (e.g., the sequence *spider-mosquito-cockroach-canary-eagle-toucan-cricket-mosquito-butterfly* would be considered as including three clusters, although *mosquito* was repeated across two clusters).

Mean cluster size: for each identified cluster (of at least three words), we counted the number of words, excluding names of a subcategory for which the participant had already generated a word and intra- (but not inter-) cluster repetitions.

Number of isolated words: the number of words of a specific semantic subcategory, not preceded nor followed by a word from the same subcategory.

Number of category switches: the number of switches was determined by the number of times the subject moved from one semantic subcategory to another, including with single words.

Number of cluster switches: number of switches from a (at least three word) cluster to another (at least three word) cluster.

Number of repetitions: any identical or variations of a previously given (and otherwise correct) word (e.g., *gato-gatinho*, meaning cat, little cat).

Number of grapho-phonological clustered responses: number of responses in which at least two successive words either began with the same phoneme(s) or letter(s), or rhymed.

Timing measures: the recorded files were digitalized and then analyzed for timing measures using *Audacity* 2.0.0 software. We first checked whether the reduction in the production rate as a function of time was independent of formal education level (as reported by [8]). To this aim we computed the number of words produced within three time slots: less than 30 s, between 30 s and one minute, and within the last minute. We then made two analyzes on interword times, namely on the time intervals between each pair of consecutive words (in seconds and hundredths). For technical reasons, the timing data of the first answers from one +FE participant were missing.

2.1.2.2. Superordinate naming. Each of the 40 trials consisted of three orally presented coordinate words (e.g., *train-plane-boat*; *canary-parrot-swallow*), and the participants were asked to tell what they had in common, or what they all were, providing only one name common to the three items. The considered near superordinate categories were: furniture, household-electric, means of transportation, combustibles, vegetables, alcoholic drinks, medias, smoke, TV actors, popular dances, metals, toys, colors, professions, flowers, cloth accessories, body parts, candies, relatives, snakes, birds, fishes, mammals, cities, countries, candidates for president, weapons, writers, condiments, insects,

cereals, planets, jewels, fruits, sports, games, disciplines, writing accessories.

The participants' performance was tape-recorded and then analyzed by five naïve judges. They categorized all participants' responses according to eight response categories: in addition to no response and clear error, the judges were asked to consider as *near superordinate* answers like *birds* to the triplet *canary-parrot-swallow*; as *distant superordinate* answers like *animal* to the same triplet, i.e., when the participant used a more generic, higher hierarchical label; as *evaluative* answers like *something that destroys people* to the triplet *cigarette-cigar-pipe*; as *syntagmatic attribute* answers like *useful for listening* to the triplet *radio-television-telephone*; as *subordinate* answers like *handle* to the triplet *sword-knife-pistol*, as *handle* is a constituent part of the three referents; and as *symbolic* answers like *death* to the triplet *sword-knife-pistol*. A response was considered as belonging to one of these categories when the majority of the judges had classified it as such. When no majority was reached, a sixth, this time non-naïve judge (one of the co-authors) chose between the most frequent categories proposed for that response by the other judges.

2.2. Results

2.2.1. Semantic fluency

We first examined the total correct output per taxonomic category and per group. A repeated-measures analysis of variance (ANOVA) by participants run on these scores showed significant effects of category,² $F(4, 116)=27.50$, $\eta_p^2=0.49$, and group, $F(1, 29)=30.65$, $\eta_p^2=0.51$, both $ps < 0.0001$. As displayed in Fig. 1, animals led to more responses than all other categories, all $p's \leq 0.01$; domestic utensils led to more responses than furniture and vehicles, $ps \leq 0.005$, and working tools to more responses than vehicles, $p < 0.001$ (all Bonferroni-corrected). The significant group by category interaction, $F(4, 116)=3.18$, $p < 0.025$, $\eta_p^2=0.10$, reflects the fact that as responses for animals were particularly numerous in the +FE participants, groups differed somewhat more for this category than for the others.³ Still, –FE participants offered less responses than +FE ones for all categories, all $ps \leq 0.005$.⁴ Finally, among the –FE participants, illiterates did not differ from semi-literates, with averages of 12.97 and 12.03 items provided, respectively.

We then compared the groups for typicality patterns. As Table 1 illustrates, there were group-dependent patterns of typicality. For example, *horse* and *dog* were the most typical animals for the –FE participants, who also mentioned *cow* in good position; in contrast, *cat* was more typical for the +FE participants, who never mentioned *cow*. For working tools, only the –FE participants answered *hoe*, and for vehicles, only the +FE participants mentioned *plane* and *boat*, whereas *truck* and *bus* (work and travel vehicles that are important for people of less advantaged classes) were mentioned in good position only by the –FE participants. After the first three responses for each category, some participants also gave responses indicating potential overlap between categories. This was the case, for example, of the attribution of *table* to both furniture and utensil by the –FE participants; category

² The category effect was also significant in the non-parametric Friedman ANOVA run on these data, $\chi^2(4)=66.28$, $p < 0.0001$.

³ Although there was no significant correlation between correct output and age (all $rs < 0.21$, $p > 0.10$), we checked that similar results were obtained when using age as covariate.

⁴ Significant group differences were also observed for all categories when applying non-parametric Mann-Whitney *U* tests to these data, $U=17, 47, 25.5, 43.5$ and 18, all $ps \leq 0.025$, for animals, working tools, furniture, domestic utensils and vehicles, respectively.

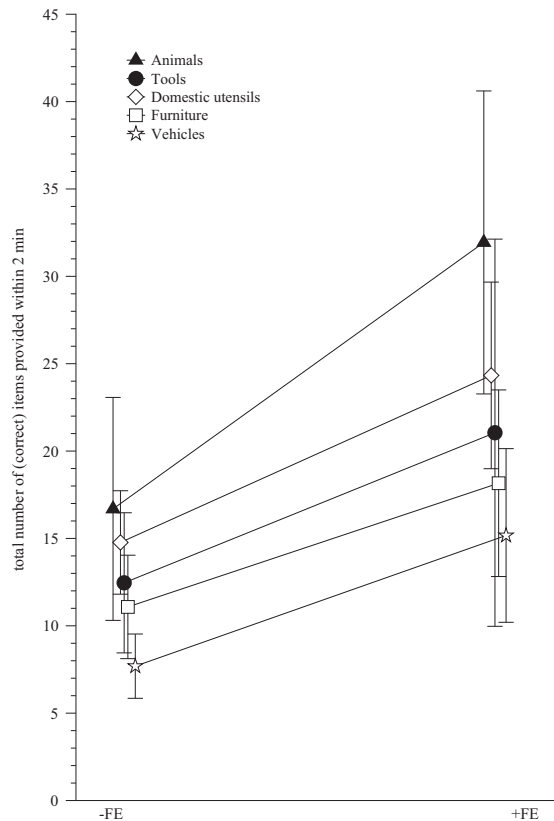


Fig. 1. Average total correct output (i.e., number of correct items provided in two minutes) in the five semantic fluency tests of Experiment 1, separately for each group (-FE: participants with null or limited formal education and +FE: participants with higher levels of formal education). Error bars correspond to standard deviations.

overlapping occurred also for *knife*, *freezer*, *stove* and *TV set*, as well as for *trailer* between working tools and vehicles.

Given that animals not only elicited the larger number of responses but were also the category that presented the least membership ambiguity and the clearest internal structure, it is only for this category that we performed further analyses. Table 2 presents the average scores, Mann–Whitney *U* tests, as well as the effect sizes and 95% confidence intervals of the differences between the means.⁵ Significant increases with level of formal education are observed for the total correct output, number of subcategories and number of clusters, and trends toward significance are observed for number of isolated words, number of between-categories switches, number of cluster switches, and number of grapho-phonologically related words (note that none of these measures differed between illiterates and semi-literates, all p s > 0.10). Thus, -FE participants did not only present a poorer total correct output than +FE participants, but also produced less subcategories, less clustered words and less isolated words, as well as less (cluster) switches and phonologically related words. Although there was a numerical trend (and a non-negligible CL effect size, see Table 2) for -FE participants to produce less words per cluster, groups did not differ strongly on mean cluster size (neither did illiterates vs. semiliterates, p > 0.10). Interestingly, cluster size only tended to be correlated with total correct output, $r(29)=0.32$, $p < 0.10$, indicating that this measure of taxonomic clustering is relatively independent of content knowledge.

⁵ As recommended by Lakens [27], to estimate effect sizes we used Hedges' *g*, which is similar to Cohen's *d*, but less biased with small samples.

Table 1

Experiment 1, semantic fluency task: items with the highest typicality score (3 or more if equality between 2) for each category and each group (-FE: participants with null or limited formal education; +FE: participants with higher levels of formal education).

	-FE	+FE
Animals		
Horse	0.41^a	0.39
Dog	0.41^a	0.41
Cow	0.26	0
Cat	0.21	0.46^a
Tools		
Hoe	0.33^a	0
Hammer	0.23	0.30^a
Knife	0.15	0.19
Saw	0	0.15
Furniture		
Sofa	0.51^a	0.44
Chair	0.38	0.63^a
Bed	0.18	0.19
Table	0.15	0.22
Domestic utensils		
Pot	0.36^a	0.09
Spoon	0.26	0.11
Table	0.15	0
Glass	0	0.15^a
Broom	0	0.11
Vehicles		
Car	0.67^a	0.35^a
Bicycle	0.18	0.17
Truck	0.18	0
Motor bike	0.15	0
Bus	0.15	0
Plane	0	0.31
Boat	0	0.25

Note: Bold scores represent the three most typical items.

^a Represents the most typical one.

No significant correlation between total correct output and number of repetitions was observed, either, $r(29)=0.29$, $p > 0.10$. On the contrary, there were significant correlations between total correct output and the six measures for which a group effect was observed, namely number of subcategories, $r(29)=0.68$, $p < 0.0001$, of clusters, $r(29)=0.80$, $p < 0.0001$, of isolated words $r(29)=0.38$, $p < 0.05$, of between-categories switches $r(29)=0.78$, $p < 0.0001$, of cluster switches, $r(29)=0.49$, $p < 0.005$, and of phonologically related words, $r(29)=0.63$, $p < 0.0001$. When statistically controlling for total output by introducing this score as covariate, no significant group difference was observed anymore on these six measures, all F 's ≤ 1 .

Hence, the main difference between -FE and +FE participants resides essentially in total output. As illustrated in Fig. 2a, this difference was particularly strong for the responses given within the first 30 s, which were particularly numerous in +FE participants. The ANOVA on correct output contrasting groups as a function of time slot (within the first 30 s, between 30 s and one minute, and within the last minute) showed significant effects of group, $F(1, 29)=28.81$, $p < 0.0001$, $\eta_p^2=0.50$, and time slot, $F(2, 58)=42.49$, $p < 0.0001$, $\eta_p^2=0.59$, as well as a significant interaction between group and time slot, $F(2, 58)=5.0$, $p < 0.01$, $\eta_p^2=0.15$.⁶ Bonferroni-corrected contrasts showed that the -FE participants presented fewer responses than the +FE participants in all time slots (all p 's < 0.05).⁷

⁶ The results were similar when age was introduced as covariate.

⁷ Consistently, Mann–Whitney *U* test were significant for the three time slots, $U=13$, $p < 0.01$, $U=38.5$, $p < 0.01$, and $U=53$, $p < 0.025$ for the first 30 s, between 30 s and one minute, and within the last minute.

Table 2
Experiment 1, animal fluency task: average scores per group (–FE: participants with null or limited formal education; +FE: participants with higher levels of formal education). Standard deviations and confidence intervals (CI) in brackets.

	–FE	+FE	Mann–Whitney (<i>U</i>)	Hedges' <i>g</i> ^a	CI ^b	CL effect size ^c
Total correct output	16.69 [6.38]	31.94 [8.67]	17**	1.90	[–20.78, –9.72]	92%
Number of subcategories	3.92 [1.11]	5.11 [1.08]	52*	1.06	[–2.01, –0.37]	78%
Number of clusters ^d	3 [1.41]	4.72 [1.49]	46.5*	1.15	[–2.80, –0.65]	80%
Mean cluster size ^d	3.68 [1.33]	4.36 [1.02]	89	0.57	[–1.58, 0.22]	65%
Number of isolated words	3.77 [1.64]	6.72 [4.88]	73.5(*)	0.74	[–5.48, –0.42]	72%
Total number of between-subcategory switches	6.77 [2.17]	14.06 [5.82]	24.5**	1.52	[–10.35, –4.23]	88%
Number of cluster switches ^d	0.77 [1.01]	1.5 [1.15]	73.5(*)	0.65	[–1.53, 0.07]	68%
Number of repetitions	1.62 [1.56]	2.50 [2.15]	87.5	0.45	[–2.24, 0.48]	63%
Number of grapho-phonological related words	2.46 [2.11]	4.28 [3.58]	71(*)	0.58	[–3.92, 0.28]	67%

^a Calculated thanks to Lakens [27].

^b 95% confidence interval of the difference between the means [low and high]; calculated thanks to Lakens [27].

^c Common language effect size, also known as the probability of superiority [27]; it reflects the chance that for a randomly selected pair of individuals, the measure is higher for +FE than for –FE; calculated thanks to Lakens [27].

^d With clusters of at least three words of the same semantic subcategory.

(*) $p < 0.10$.

* $p < 0.025$.

** $p < 0.01$.

The analysis ran on interword times suggest in addition that content knowledge is more rapidly exhausted in –FE than in +FE participants. Indeed, as shown in Fig. 2b, when interword times were analyzed as a function of retrieval position (as in [36], i.e., independently of the fact that responses were intra- or inter-category), the slope of the retrieval-position function estimated for the first 16 positions is steeper in –FE than in +FE participants. As the 16th position corresponds to the average number of answers given by the –FE participants, whereas it is only about half of the average number of responses provided by the +FE participants, for –FE participants search for new exemplars would become sooner more difficult and hence slower in the retrieval-position function, compared to +FE participants. The results of a reanalysis of the data using relative scores are coherent with this interpretation. In this reanalysis, we calculated for each participant individually his/her interword times for his/her first, second and third thirds of the responses (i.e., on the first three, middle three and last three items for a participant who gave only 10 answers, hence for whom nine time intervals were calculated, but on the first 10, middle 10 and last 10 items for a participant who provided a total of 31 responses). As Fig. 2c illustrates, no slope difference is observed anymore. This was confirmed by an ANOVA with group and retrieval position (first, middle or last third of the response positions, as defined individually) as factors: only the main effects of group, $F(1, 28)=5.77$, $p < 0.025$, $\eta_p^2=0.17$, and of retrieval position, $F(2, 56)=27.18$, $p < 0.001$, $\eta_p^2=0.49$, were significant, not the interaction between these two factors, $F < 1$, $\eta_p^2=0.01$.⁸

However, as illustrated in Figs. 2b and c, –FE participants presented somewhat longer interword times than +FE participants overall ($M=4.12$ s, $SD=2.29$, and $M=2.75$ s, $SD=1.10$, respectively). Yet, the difference between groups was not significant, Mann–Whitney $U=170$, $p > 0.10$. This is probably due to the

fact that –FE participants also tended to stop answering slightly before the +FE participants (total average production time: $M=1$ min and 41 s, $SD=32.3$ s, and $M=1$ min and 52 s, $SD=15.74$, respectively). Yet, this effect was not significant, either, Mann–Whitney $U=103.5$, $p > 0.10$. In fact, –FE participants were quite variable, some stopping answering before 30 s, and others going until the end of the allowed time slot.

Critically, interword times were also calculated separately for intra- and inter-subcategory consecutive pairs of responses. Since according to the Wilcoxon matched pairs test there was no difference between cluster switches and hard switches, $Z=0.23$, $p > 0.10$, we contrasted all inter-categories responses to intra-subcategory responses in an ANOVA taking group and interval type (inter- vs. intra-subcategory) as factors. There were significant main effects of group, $F(1, 29)=5.59$, $p=0.025$, $\eta_p^2=0.16$, and of interval type, $F(1, 29)=52.18$, $p < 0.0001$, $\eta_p^2=0.64$, with longer interword times in –FE than +FE participants, and for inter- than intra subcategory responses (on the average, 5.15 vs. 1.92 s, respectively⁹). However, as illustrated in Fig. 2d, the interaction between these variables was not significant, $F(1, 29)=2.10$, $p > 0.10$, $\eta_p^2=0.03$.¹⁰

2.2.2. Superordinate naming

As displayed in Fig. 3, –FE participants presented a lower proportion of near superordinate responses than +FE participants, Mann–Whitney $U=27$, $p < 0.0025$,¹¹ (with almost no difference between illiterates and semi-literates: 58.03% vs. 59.29%, respectively).

⁹ This difference was also significant according to the Wilcoxon matched pairs test, $Z=4.89$, $p < 0.0001$.

¹⁰ The results were similar when age was introduced as covariate. In addition, the difference between groups was not significant in the Mann–Whitney U test ran on the difference between inter- and intra-category times, $U=152$, $p > 0.10$, which corresponds to the group by interval type interaction in the ANOVA.

¹¹ A parametric analysis with age as covariate led to a similar result, with a significant group effect, $F(1, 28)=15.91$, $p < 0.0005$, $\eta_p^2=0.36$.

⁸ The results were similar when age was introduced as covariate. In addition, the difference between groups was not significant in the Mann–Whitney U test ran on the difference between the first and last third of retrieval positions, $U=101$, $p > 0.10$, which corresponds to the group by retrieval position interaction in the ANOVA.

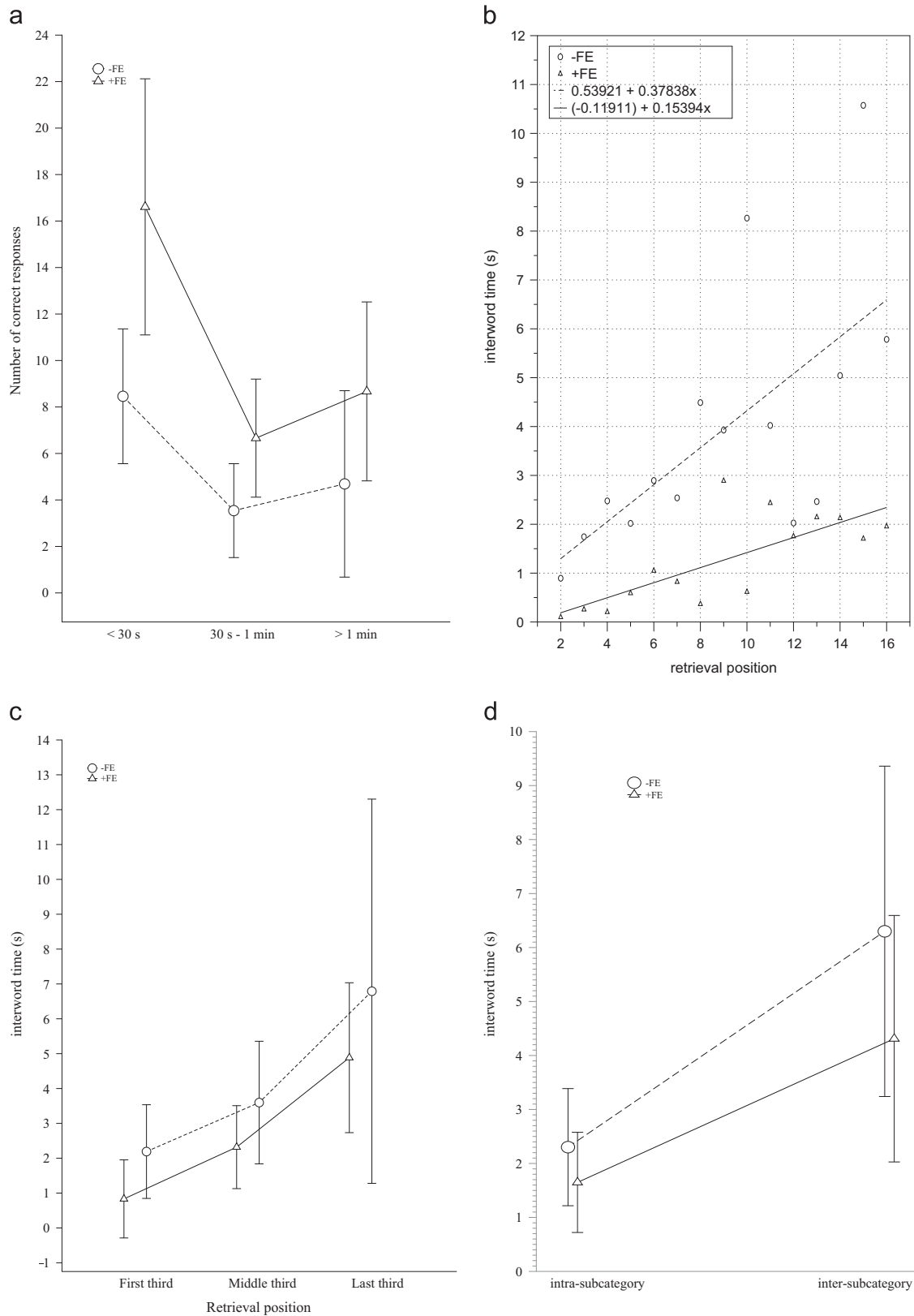


Fig. 2. Experiment 1, animal fluency test, average results, separately for each group (–FE: participants with null or limited formal education and +FE: participants with higher levels of formal education). Error bars correspond to standard deviations. (a) Correct responses produced within three time slots: within the first 30 s, between 30 s and one minute, and within the last minute. (b) Interword times as a function of retrieval position; also shown are the best fitting linear regression lines and corresponding parameters. (c) Interword times as a function of relative retrieval position (defined individually as responses given in the first, second and third thirds of the responses); also shown are the best fitting linear regression lines and corresponding parameters. (d) Interword times for consecutive responses belonging to either the same semantic subcategory (intra- subcategory) or to different semantic subcategories (inter-subcategory).

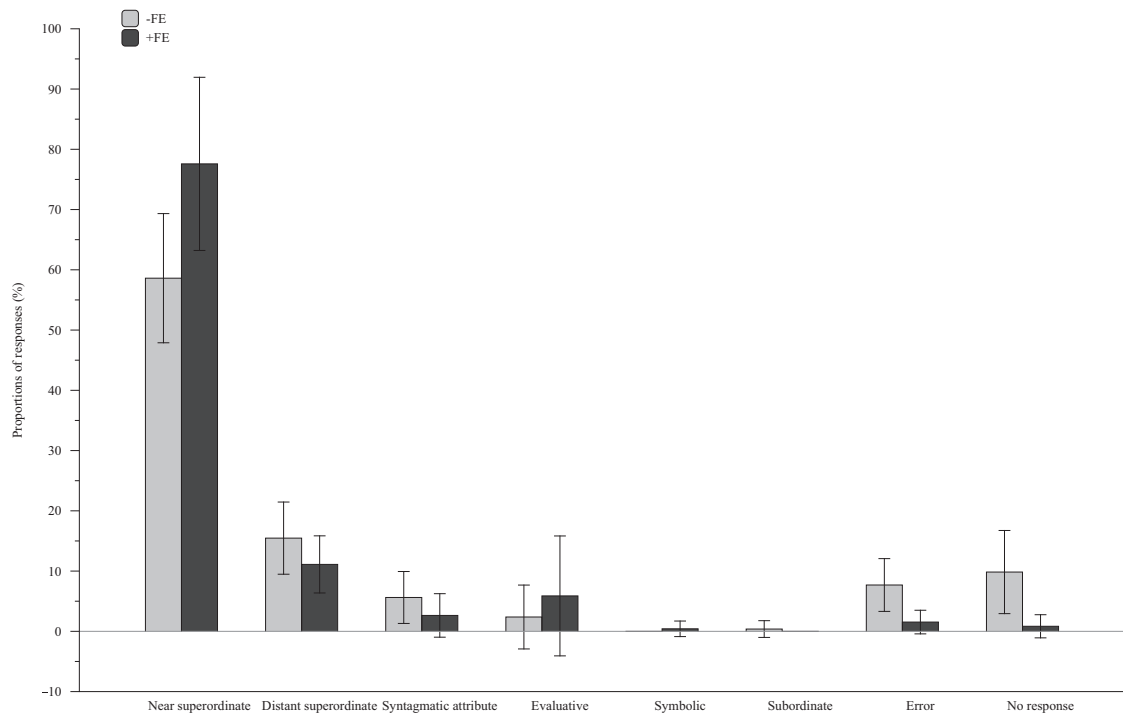


Fig. 3. Experiment 1, superordinate naming task: proportion (in %) of the different types of responses, separately for each group (–FE: participants with null or limited formal education and +FE: participants with higher levels of formal education). Error bars correspond to standard deviations.

Qualitative analysis of the –FE participants' wrong answers or absence of response showed that they did not know the near superordinate name or concept. For example, none of the –FE but 89% of the +FE participants were able to say that *copper*, *iron*, and *zinc* are metals. This was probably due, in most –FE participants, to either a lack of knowledge of the name *metal* or to difficulties in rapidly accessing to it. In contrast, the fact that, for example, only half of the +FE participants (no illiterate and 17% of the semi-literates) were able to say that *bat*, *dolphin* and *whale* are mammals was more likely due to either a difficulty of access to the mammal category given the low typicality of those instances, or to a taxonomic confounding of any of these concepts with the bird or fish categories (e.g., commenting on the *dolphin* and *whale* stimuli, an illiterate said that they were *ocean fishes*, of course). Lack of knowledge of the instances rather than of the category may have been at the origin of the fact that only 57% of the illiterates, 67% of the semi-literates, but 89% of the +FE participants were able to say that *Goiânia*, *Palmas* and *Natal* are three Brazilian cities.

We next calculated the correlations between the proportion of subordinate responses in the superordinate naming task and each of the measures we had calculated in the animal fluency test. Notably, no correlation was significant (all p 's > 0.10), except with total correct output, $r(29)=0.43$, $p < 0.025$. Also, interestingly, the correlation with correct output in the other fluency tests was significant for furniture and vehicles, $r(29)=0.54$ and 0.52 , respectively, both p 's < 0.01, but only tended towards significance for domestic utensils, $r(29)=0.31$, $p < 0.10$, and was not significant for working tools $r(29)=0.28$, $p > 0.10$.

2.3. Discussion

The results of the present experiment confirm the hypothesis that formal education mainly impacts on content knowledge. In the fluency task, –FE participants produced less correct total output, less subcategories, less clustered words and less isolated words, as well as less (cluster) switches and phonologically related

words than +FE participants. Consistently, in the superordinate naming task, –FE participants presented less near superordinate responses than +FE participants, which reveals lack of detailed and precise content knowledge, in particular at the lexical level. Furthermore, the typicality patterns observed in the various fluency tests also differed between –FE and +FE participants, reflecting different experiences. This is consistent with the observations made by Brucki and Rocha [8], who also reported that the illiterate unschooled people they examined did not mention the same prototypical animal as the other groups. In particular, as in the present animal fluency test, they frequently mentioned the horse, while the most educated participants of their study mentioned frequently the cat.

Yet, in the animal fluency test the two groups of participants did not differ on cluster size, and the group differences observed on the six measures other than total output were canceled once correcting for individual differences in total correct output through covariate analyzes. In addition, in animal fluency, although –FE participants tended to be slower to provide their answers, with somewhat longer interword times than +FE participants, they were so overall, for both intra- and inter-category responses. This suggests that for –FE participants both the duration of each switch and the duration of each retrieval were prolonged. Importantly, they presented shorter intra- than inter-subcategory responses, as did the +FE participants. This confirms retrieval by subcategory in both groups. Thus, independently of content knowledge, neither taxonomic clustering, nor the basic mechanisms of access to subcategories seem to depend on formal education level. In addition, although the analysis on interword time intervals as a function of retrieval position suggests a more rapid exhaustion of content knowledge in –FE compared to +FE participants, no special intrinsic difficulty of –FE participants in accessing semantic information is observed when retrieval positions are defined relatively, as a function of total correct output.

Nevertheless, which factor of formal education is responsible for the effects observed on content knowledge remains unclear from the present data. Indeed, in the present experiment literacy

was confounded with formal education, in terms of schooling years. In addition, although no clear difference between illiterates and semi-literates was observed, we must acknowledge that these subgroups only included seven and six participants, respectively. Hence, no definite conclusion can be drawn as regards the possible specific impact of literacy on the various measures of fluency. Experiment 2 was aimed at examining this question more thoroughly.

3. Experiment 2: Semantic fluency and literacy

To examine specifically the relation between reading proficiency and semantic fluency, here we examined participants who were all almost unschooled but quite heterogeneous as regards their literacy level, which was carefully evaluated through various reading-related measures. More precisely, we examined 28 unschooled adults displaying varying degrees of rudimentary reading decoding, most of them attending the first class of a literacy program for adults.

They were presented with two different tests of semantic fluency. The first was an animal fluency task similar to that of Experiment 1. The second was an alternating fluency task in which participants had to produce in alternation the name of a fruit and the name of a tool. Thus, the latter task imposes comparable demands upon semantic memory as a simple semantic fluency task (for both, participants are requested to find words belonging to a specified semantic category), and is equally constrained by speed (here, one minute for each test), but requires constant switching between concepts, which probably involves greater executive demands compared to conventional fluency tasks (e.g., [12,19,20,36,67]). It thus allowed us to further investigate the impact of literacy on the switching component of fluency performance. Furthermore, we used as a control two tests of phonemic fluency, in which participants were required to give words beginning with a given letter or phoneme. Indeed, the effect of education on overall performance is much stronger in phonemic compared to semantic fluency [23,48], probably because phonemic fluency involves phonological awareness, which is strongly linked to reading ability [41].

In addition to these fluency tasks, reading-related abilities were evaluated through identification of letter and graphemes, reading tests (word and pseudoword reading, plus written text comprehension), and metaphonological tests that required manipulating either syllables (inversion and deletion tests) or phonemes (inversion and deletion tests, plus production of acronyms). Using the *Standard Progressive Matrices (PM38)* [49], participants were also evaluated as regards their analogical reasoning capacity. Finally, as working memory seems important for verbal fluency performance due to the need to self-generate category cues and monitor the products of retrieval [54,71], we also presented them with two memory tests. The critical one was *listening span*, a working memory test in which participants have to listen to a series of sentences and retain the final word of each sentence for spoken recall at the end of the series while simultaneously judging each sentence as true or false [65]. An additional test of *word span* controlled for phonological memory.

3.1. Method

3.1.1. Participants

Among the 28 participants, 22 were attending the first class of a literacy course for adults in the quarter of Capoeiras, Florianópolis, Santa Catarina, Brazil. The other six were living in the same area. They were aged from 17 to 77 years (average: 37.5 years; 15 women). Most of them (15) had never attended school; some had

attended school for some months (5 participants) or even a year (6 participants) or more (18 months: 2 participants), but in an irregular way (average: 4.7 months). They were rewarded for participation by receiving school material.

All participants were screened for reading-related knowledge through the following 10 tests [59]: letter identification (23 letters), grapheme identification (21 graphemes such as *ã*, *im*, *lha*), word reading (20 trials, all with simple frequent words such as *chuva* – meaning “rain”, *feira* – meaning “feast”, *tigela* – meaning “bowl”), pseudoword reading (20 trials, all derived from the words presented in the former test, e.g., *cuda*, *vesta*, and *figeta*), reading comprehension (11 questions about a short, 89 words long text, called *O tatu encabulado*, meaning “The embarrassed armadillo”), oral syllable deletion (16 CVCV trials), oral syllable reversal (10 CVCV trials), phoneme deletion (26 trials: 16 CVC and 10 CCV), phoneme reversal (20 trials: 5 CV, 5 VC, 10 VCV) – note that most stimuli and responses were pseudowords in the latter four tests – and auditory acronyms (16 trials, e.g., *charmosa ilha* – meaning “charming island”). A total literacy score was calculated as the average of performance on these 10 tests. As can be seen in Table 3, although reading-related performance was poor in most participants, it was also highly variable.

Participants also completed the five series (60 items) of the PM38 [49]. Scores on the PM38 were generally low and with a considerable dispersion. This is not surprising, as a recent Brazilian study showed large national discrepancies, with limited education level and low socioeconomic status deeply affecting PM38 performance [15]. In addition, scores as low as 12.5 are commonly observed in illiterate and/or unschooled participants [73].

Participants were also presented with two memory tests. In the word span test, they had to repeat sequences of increasing length, each sequence including two to five monosyllabic frequent Portuguese words (*chão*, *pé*, *mar*, *céu*, *gol*, meaning “ground”, “foot”, “sea”, “sky”, “goal”). Seven different sequences were presented for each sequence length. In the listening span task, participants had to listen to a series of sentences, retaining the final word of each sentence for spoken recall at the end of the series, while simultaneously judging each sentence as true or false. Here, to control for word familiarity, all sentences ended with a word denoting a numeral. After an initial example including two sentences (*Meia dúzia são seis*; *Os dias da semana são sete*, meaning “Half a dozen is six; the days of the week are seven”); expected answer: *seis* – *sete*, “six-seven”), series of three to five sentences were presented. Seven different trials (series of sentences) were presented for each

Table 3

Experiment 2: mean age and schooling level of the participants. Correct scores (in %) observed in the reading-related tests and performance in the PM38. SD=standard deviations. Min and Max: minimum and maximum observed scores, respectively.

	Mean	SD	Min	Max
Age (yrs)	37.54	14.37	17	77
Schooling (months)	4.73	6.1	0	18
Letter knowledge (%)	77.34	28.75	4	100
Grapheme knowledge (%)	46.34	30.02	0	100
Word reading (%)	45.63	37.44	0	100
Pseudo-word reading (%)	32.14	30.14	0	90
Written text comprehension (%)	42.37	38.77	0	100
Phoneme deletion (%)	33.38	38.68	0	100
Phoneme reversal (%)	14.91	22.57	0	70
Acronyms (%)	11.83	26.43	0	88
Syllable deletion (%)	53.46	28.65	0	100
Syllable reversal (%)	31.79	34.86	0	100
Total literacy score (%)	38.92	26.86	2.93	89
Standard progressive matrices (PM38)	16.54	4.56	9	29
Listening span	3.09	0.78	2	4.5
Word span	2.96	0.70	1.5	4

length. For both tests, participants were presented with sequences of increasing length, proceeding to the next length only if they had answered correctly to at least four of the seven trials. For each test, the span was estimated on each participant as being the length for which the participant presented at least four correct answers; this score was then corrected as a function of the number of correct trials obtained for the next length (with 0.25, 0.50 and 0.75 scores for 1, 2 or 3 correctly answered trials, respectively). Average results are presented in Table 3.

3.1.2. Fluency tasks: materials, procedure and scoring

In the phonemic fluency task, participants were asked to produce for one minute as many words as possible starting with either /a/ or /p/. In the animal fluency task, participants were asked to produce for one minute as many animal names as possible, without repetition. In the alternating semantic fluency task, they had to produce for one minute alternatively the name of a fruit and of a tool, also without repetition. For all tests, the experimenter used a chronometer and stopped the participant after one minute, tape-recording the session for further data analysis. For the animal fluency task, the scoring was the same as in Experiment 1, except that no digitalized version of the tape-recorded data could be provided. Timing parameters were thus estimated on the digital recordings, using a chronometer. The same holds true for the alternating fluency task. For the latter, we also estimated alternation errors, for instance answering with two tool names in a row.

All participants were first presented with the phonemic fluency tests (first with /a/ then with /p/), next with animal fluency, and finally with alternating fluency. The reading-related tests were presented before the fluency tasks, the PM38 afterwards. All tests were parts of a larger battery of 56 tests that were presented to participants in eight to 10 sessions over a period of six months [38].

3.2. Results

Table 4 presents the average values of the various measures of fluency performance in the three fluency tasks, as well as the correlations between these scores and age, schooling (in months), literacy (total literacy score), and performance in the PM38 and memory tests. For the scores that significantly correlate with literacy, Table 4 also presents the partial correlations that were computed when a measure tended to correlate with total output, PM38, working memory, or word span (this was done when $r(28) \geq 0.25$).¹²

As shown in Table 4, the average total correct output was higher in the animal fluency task than in the other two tasks, with the lowest performance observed in the phonemic fluency task (on the average, 4.82 and 6.14 total correct output for /a/ and /p/, respectively). Consistently, the ANOVA run on these scores showed a highly significant effect of task,¹³ $F(2, 54) = 42.47$, $p < 0.0001$, $\eta_p^2 = 0.61$ with better performance in the animal than in both the alternate, $F(1, 54) = 9.55$, $p < 0.005$, and phonemic fluency tasks, $F(1, 54) = 82.16$, $p < 0.0001$, and in the alternate than in the phonemic fluency task, $F(1, 54) = 35.7$, $p < 0.0001$. Yet, total correct output was strongly correlated between these three tasks (between animal fluency and alternating fluency: $r(26) = 0.62$, $p < 0.0005$; between animal and phonemic fluency: $r(26) = 0.70$,

$p < 0.0001$; between phonemic and alternating fluency: $r(26) = 0.48$, $p < 0.01$).

As illustrated in Table 4, there was no significant correlation at all between the various fluency scores and age or number of months during which the participants attended school in childhood, probably because many of them (54%) never attended school. Total correct output did correlate with literacy in all three tasks, and remained significantly correlated with literacy after removing the effect of either working memory or PM38 in both animal and phonemic fluency, but not in alternating fluency, in which the correlation vanished after removing the effect of working memory. As can be seen in Table 4, all the correlations of the other fluency measures with literacy vanished after removing the effect of either working memory or PM38.

As in Experiment 1, we examined interword times in two ways. First, we analyzed them as a function of retrieval position, contrasting the animal to the alternating fluency tasks. As shown in Fig. 4, when intervals were computed over the first nine positions (which corresponds to the average total correct output in alternating fluency), the slope of the retrieval-position function is somewhat steeper in the alternating than in the animal fluency task. This confirms the idea that alternating fluency is more difficult than animal fluency, probably because it involves higher executive functioning (and in particular working memory) demands (e.g., [12,19,20,36,77]).

Second, as in Experiment 1 we also examined whether interword times were shorter when these belonged to the same semantic subcategory than when they belonged to different subcategories. As there was no strong difference in interword times between cluster switches and hard switches (with averages of 5.91 and 6.29 s, respectively), and as eight participants did not present any cluster switch, we contrasted all inter-subcategories responses to intra-subcategory responses (see Table 4). The Wilcoxon matched pairs test confirmed that interword times were significantly longer for inter- than intra subcategory responses, $Z = 2.25$, $p < 0.025$.

In addition, it is worth noting that the interword times in alternating fluency, which by definition were inter-category, are almost identical to, and in any case not longer than, the inter-category interword times of the animal fluency task (see Table 4). Yet, these measures did not correlate with each other, $r(26) = 0.22$, $p > 0.10$ (neither did the interword times of alternating fluency with the intra-category interword times of animal fluency, $r(26) = 0.10$, $p > 0.10$).

3.3. Discussion

The present experiment clearly shows that literacy, even of a rudimentary level, contributes to acquire content knowledge: in animal fluency, the total number of generated words strongly correlated with literacy level, and this correlation cannot be explained by either working memory or general intelligence, as estimated through the PM38. The same holds true for total output in phonemic fluency, but not for total output in alternating fluency. Indeed, although total output in the alternating fluency task was marginally correlated with literacy, this correlation was not significant any longer after partialing out the effect of working memory. Together with the fact that the number of (cluster) switches did not correlate with literacy in animal fluency, this suggests that the switching component of semantic fluency, be it spontaneous as in the simple animal fluency task, or imposed as in the alternating fluency task, does not strongly depend on literacy. Similarly, the fact that the correlations of the other fluency measures with literacy vanished after removing the effect of either working memory or PM38 suggests that literacy per se affects only content knowledge, more particularly total correct output. In

¹² Note that the partial correlations that remove the effect of total output, PM38, working memory, or word span also take into account the correlation between literacy and total output, PM38, $r(26) = 0.54$, $p < 0.01$, working memory, $r(26) = 0.58$, $p < 0.01$, or word span, $r(26) = 0.14$, $p > 0.10$, respectively.

¹³ The task effect was also significant in the non-parametric Friedman ANOVA run on these data, $\chi^2(4) = 33.82$, $p < 0.0001$.

Table 4

Experiment 2: average scores in the animal and alternating semantic fluency tasks, as well as in phonemic fluency. SD: standard deviations. WM: working memory.

Fluency task	Mean	SD	Correlations											
			Literacy _{total} output	Schooling	PM38	Total output	Age	WM	Word span	Lit _{total} output	Literacy _{schooling}	Literacy _{PM38}	Literacy _{WM}	Literacy _{Word} span
Animal														
Total correct output	11.04	3.95	0.57***	-0.14	0.39*	-	0.05	0.48**	0.15	-	-	0.46*	0.46*	-
Number of subcategories	3.07	0.86	0.19	-0.21	0.09	0.49***	-0.04	0.29	-0.12	-	-	-	-	-
Number of clusters ^a	1.75	1.04	0.39*	0	0.26	0.74***	0.10	0.27	0.15	-0.06	-	0.31	0.30	-
Mean cluster size ^a	3.55	1.55	0.40*	0.29	0.36 ^(*)	0.26	0.02	0.09	0.18	0.32 ^(*)	0.35 ^(*)	0.26	-	-
Number of isolated words	2.29	2.48	-0.07	-0.20	-0.02	0.26	-0.04	0.20	-0.05	-	-	-	-	-
Total number of between-subcategory switches	4.54	2.6	0.19	-0.16	0.08	0.51***	0.14	0.39*	-0.04	-	-	-	-	-
Number of cluster switches ^a	0.29	0.54	0.03	-0.02	-0.10	0.24	0.11	-0.06	0.01	-	-	-	-	-
Number of repetitions	0.54	0.69	0.12	0.04	0.12	0.20	0.06	0.15	-0.10	-	-	-	-	-
Number of grapho-phonological related words	0.89	0.83	0.37*	-0.19	0.52***	0.64***	-0.05	0.41*	0.25	0.01	0.44**	0.12	0.18	0.35 ^(*)
Cluster switches interword time	5.91	5.34	-0.33 ^(*)	0.24	-0.25	-0.50***	0.09	-0.40*	0.14	-0.06	-0.42*	-0.24	-0.13	-
All inter-category interword time	6.29	3.64	-0.28	0.08	-0.14	-0.38*	0.26	-0.19	0.03	-	-	-	-	-
Intra-category interword time	4.43	3.05	-0.38*	0.04	-0.30	-0.54***	0.04	-0.49***	-0.05	-0.10	-	-0.27	-0.14	-
Normalized switching time cost ^b	18.95	32.20	0.08	-0.07	0.08	0.20	0.12	0.34 ^(*)	-0.05	-	-	-	-	-
Alternating														
Total correct output	9.14	3.46	0.37*	0.10	0.07	-	0.08	0.62***	0.46**	-	-	-	0.02	0.35 ^(*)
Number of repetitions	0.43	0.79	0.04	-0.11	-0.02	-0.04	0.30	-0.05	-0.02	-	-	-	-	-
Number of alternation errors	0.57	1.20	-0.16	0.08	0.30	-0.51***	0.02	-0.08	-0.02	-	-	-	-	-
Number of grapho-phonological related words	1.14	1.24	0.02	-0.13	0.06	0.13	-0.19	0.17	0.21	-	-	-	-	-
Interword time	6.08	3.80	-0.27	-0.15	-0.16	-0.56***	-0.23	-0.53***	-0.18	-	-	-	-	-
Phonemic														
Total correct output (averaged for /a/ and /p/)	5.48	3.46	0.87***	0.07	0.49***	-	-0.07	0.61***	0.15	-	-	0.83***	0.80***	-

^a With clusters of at least three words of the same semantic subcategory.

^b The difference between inter- and intra-category interword times, normalized as a function of overall speed.

(*) $p \leq 0.10$.

* $p \leq 0.05$.

** $p \leq 0.025$.

*** $p \leq 0.01$.

**** $p \leq 0.001$.

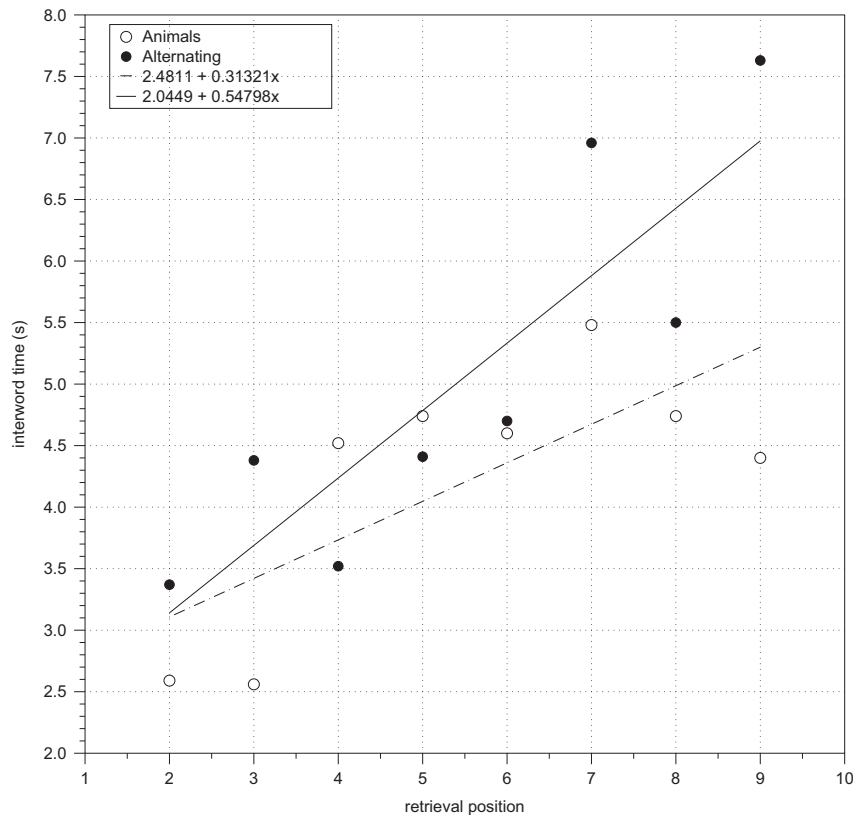


Fig. 4. Average interword times in Experiment 2 as a function of task (animal vs. alternating fluency); also shown are the best fitting linear regression lines and corresponding parameters.

addition, in animal fluency, participants presented shorter intra- than inter-subcategory interword times, showing retrieval by subcategory.

4. General discussion

Previous work has suggested that formal education in the Western culture discourages thematic (e.g., functional) thinking (e.g., [30,60,61]) and enhances the richness and precision of semantic knowledge (i.e., content knowledge), with better semantic fluency performance (i.e., number of words generated for a specific taxonomic category) in formally educated adults than in adults having no formal education [8,10,17,23,48,50,62]. Yet, less attention has been paid to the influence of formal education, including literacy, on the structure of semantic categories and on the mechanisms used to access them. In addition, in all former studies on semantic fluency, formal education and literacy levels were confounded, as they contrasted illiterate, unschooled people to formally educated, literate people.

In the present study, we examined the hypothesis that, contrary to content knowledge, the acquisition of taxonomic knowledge and the basic access processes to that knowledge do not depend on formal education and literacy. By this view, even people with limited formal education (including totally illiterate people) would represent taxonomic relations and categories, as neither formal education nor literacy by itself change the way entities are represented in conceptual memory, including their hierarchical taxonomic organization, nor the basic access processes to that information.

To examine this hypothesis, we took advantage of the fact that the semantic fluency task may reflect not only content knowledge, but also the way entities are represented in conceptual memory,

including their hierarchical taxonomic organization. We analyzed the semantic fluency data collected in participants of various levels of formal education (Experiment 1) or literacy (Experiment 2) not only in terms of overall performance but also in terms of the way they structured their responses and the time they needed to do it. More precisely, in Experiment 1, we contrasted two groups of participants that differed by their level of formal education: either with up to four years, or between 11 and 14 years of formal education. In Experiment 2, we examined more specifically the relation between reading proficiency and the various facets of semantic fluency on a fresh group of participants who had never or almost never attended school in childhood but presented varying degrees of rudimentary reading decoding, as most of them were attending the first class in an adult literacy course.

The influence of formal education on content knowledge was confirmed by the data of the semantic fluency tasks of both Experiments 1 and 2. In Experiment 1, in which we contrasted participants with limited or higher levels of formal education, this influence was observed on two measures of semantic fluency. First, across the five fluency tests we used (animals, working tools, furniture, domestic utensils, vehicles), the participants with higher levels of formal education generated more words per taxonomic category than those with limited formal education. Second, for animals (the category that we analyzed more thoroughly), the participants with higher levels of formal education also produced a greater number of different semantic subcategories.

Interestingly, across the five tests, performance varied as a function of category, with a stronger group effect for animals than for domestic utensils and working tools, confirming the importance of the ecological or cultural relevance of the chosen semantic criteria as well as of the level of reference to concrete knowledge and specific situations [17,51,57]. This was confirmed

by qualitative between-groups differences in terms of typicality patterns, consistent with Brucki and Rocha's [8] observations, and which probably reflect different experiences. Specifically, fewer pieces of furniture are likely to be found in the homes of the less formally educated participants, who also have a lower socio-economic status than the ones with higher levels of formal education. Similarly, most of the less formally educated participants probably never traveled by plane or boat. As regards animals, as the present participants are urban populations, it seems likely that the number of animals generated is tied to awareness of names of exotic species (e.g., lion, kangaroo), i.e., to encyclopedic knowledge. Nevertheless, the influence of formal education on total output was observed here even with "ecological" categories like domestic utensils and working tools (see [62] for similar data).

In accordance with these fluency data, in the superordinate naming task used in Experiment 1, in which participants had to provide a superordinate name (e.g., *furniture*) for three orally presented coordinate words (e.g., *sofa-table-cupboard*), less formally educated participants presented fewer near superordinate responses than more formally educated participants. Limited formal education thus seems to be associated with a lack of precise taxonomic knowledge and associated verbal labels. In addition, we observed significant correlations between the proportion of subordinate responses and the total number of words generated in the fluency task for the categories of animals, furniture and vehicles, but not for working tools and domestic utensils. This reinforces the idea that the total output in semantic fluency reflects content knowledge, which varies as a function of the ecological or cultural relevance of the semantic criteria.

In Experiment 2, we showed that part of the effect of formal education on content knowledge is related to literacy level. As a matter of fact, in Experiment 2 we presented the animal semantic fluency task to adults with limited formal education displaying various degrees of rudimentary literacy, which level was carefully evaluated. The total number of animal names generated (but, somewhat surprisingly, not the number of subcategories) correlated strongly with literacy level, an effect that cannot be explained by either working memory or general intelligence, as estimated through the PM38. Thus, the richness of taxonomic knowledge seems affected by both literacy and formal education.

However, other aspects of the data of the animal fluency task used in Experiments 1 and 2 showed that illiterates and/or people with limited formal education do represent and access taxonomic categories in the same way as literates and/or people with higher level of formal education. Indeed, in the animal fluency task of Experiment 1, cluster size, which indicates taxonomic clustering relatively independently of content knowledge, did not differ between participants with limited vs. high levels of formal education. In the similar animal fluency task of Experiment 2, although cluster size did correlate with literacy, this effect was canceled by partialing out the effect of general intelligence, as estimated through the PM38. Furthermore, in the animal fluency tasks of both Experiments 1 and 2, participants with limited formal education, including fully illiterates, presented shorter interword times for intra- than for inter-subcategory responses, as did the participants with higher levels of formal education of Experiment 1. This confirms that retrieval by subcategory is independent of formal education and literacy level. Thus, independently of content knowledge, even people with limited formal education, including fully illiterate subjects, display taxonomic clustering and retrieval by subcategory.

Yet, formal education and literacy slightly speed the search for new exemplars. Indeed, in the animal fluency task of Experiment 1, participants with limited formal education tended to be slower overall in finding exemplars than participants with

higher levels of formal education. In addition, in the animal fluency task of Experiment 2, the intra-category interword time was negatively correlated with literacy, and the negative correlation between literacy and cluster switches interword time tended towards significance. This suggests that education and literacy speed exemplar retrieval. The reasons for this may be twofold. On the one hand, literacy and formal education may have a generic effect: people with low levels of formal education may have general difficulties at speeded responses, to which they are not used (e.g., [72]). On the other hand, for people with higher levels of education, who are literate, search for new exemplars may be speeded by strategically using cues that are less available to illiterates and/or people with limited education, as grapho-phonological cues. In agreement with this idea, in the animal fluency task of both Experiments 1 and 2, the number of grapho-phonological related words was highly correlated to total output. In Experiment 1, grapho-phonologically related words were more numerous in participants with higher levels of formal education, and in Experiment 2 there was a significant correlation between the use of grapho-phonological cues and literacy level. This thus suggests that formal education, and, in particular, literacy does not only allow gaining new semantic knowledge, but also affords new retrieval cues. Yet, the use of such cues seems also strongly linked to general intelligence and working memory, and in Experiment 2 the correlation between literacy and number of grapho-phonological related words was canceled when partialing out these factors.

Note that the effects of literacy were observed even though the literate participants examined in Experiment 2 had actually a rudimentary level of literacy. They were most likely stimulated by various reading-related activities proposed in their literacy class, for instance to elaborate sentences including names of fruits, to write down shopping lists, to write names of animals, people, food, drinks, clothes, etc., to read and copy texts, and to detect or complete rhyming words [33]. Thus, both formal education and reading-related activities, at least when these are, in fact, more general educative activities, facilitate access to informational elements that people cannot access through direct experience and hence offer the opportunity to broaden and refine knowledge about different semantic categories.

A further component of access to semantic information is the ability to shift to another subcategory after exhausting the previous one, which has been related to executive functions and frontal lobe functioning [69,70]. This was estimated by examining the number of spontaneous between-subcategory switches in the animal fluency task of Experiments 1 and 2. There was an influence of formal education on the number of switches (Experiment 1), but no significant link with literacy level (Experiment 2). In Experiment 2, we further investigated the impact of literacy on the switching component of fluency performance by using an alternating fluency task, in which switching was imposed and continuous, as participants had to produce in alternation the name of a fruit and the name of a tool. This continuous switching task is known to induce greater difficulties in aged people or neurological patients compared to classic fluency tasks (e.g., [20,36,77]). This was observed here too: total correct output was lower, and the slope of the retrieval-position function was steeper in alternating compared to single-category (animal) fluency. More importantly for our purpose, although literacy did correlate with total correct output in the alternating fluency task, this correlation was not robust. Indeed, it did not resist partialing out the correlation with working memory. Furthermore, across the two semantic fluency tasks used in Experiment 2, the measures of inter-subcategories switching times were not correlated to each other, suggesting that they tap different facets of mental flexibility. In fact, the switching cost in the alternating fluency task may rather reflect task-set

inhibition, namely the capacity to select the appropriate task-set from competing possibilities (e.g., [2]).

Work on executive processes in illiterate and/or unschooled people is scarce, and most of it concerns language and offers relatively indirect evidence. For instance, Petersson and colleagues [44,45] applied a network approach based on structural equation modeling on PET data collected in word and pseudo-word repetition [9]. They reported a difference between literate and illiterate adults that they attributed to executive aspects of verbal working memory, as subserved by the left middle-inferior dorsolateral prefrontal cortex. Here, in Experiment 2 we also observed a strong relation between literacy and working memory, as measured through a listening span task.

Yet, executive functions are not monolithic: they form a set of related but distinct functions (e.g., [16,39]). Knowing what exactly are the “working” aspects of the so-called working memory tasks, whether they are specifically dedicated to working memory or “borrowed” from other cognitive functions and where they are located in the brain remain a matter of debate (see reviews in e.g., [5,11]). Hence, although an effect of literacy has been observed on working memory performance [25], the specificity of the link with the switching component in semantic fluency as well as with other executive processes remains obscure. As illustrated by the strong correlations with working memory observed in the animal fluency task of Experiment 2 for both total correct output and speed of intra-category access, executive processes, among which working memory, are probably relevant not only for switching but also for every single act of retrieval, including updating the current search criterion and stopping or initiating single retrieval processes [35,36].

In any case, the present switching data are consistent with other results briefly mentioned by Morais and Kolinsky [40]. In an adaptation of the California Card Sorting Test (which implied shifting between five alternative classification criteria), participants with high levels of formal education found more criteria, but no shifting difference was observed between illiterates, semi-literates and late literates (who never attended school in childhood but learned to read at adult age).

Our study further suggests that another component of executive processes that is involved in fluency tasks, namely the capacity to monitor the products of retrieval to prevent errors [54,71], is also unaffected by formal education and literacy. As a matter of fact, in Experiment 1 no group difference was observed in terms of the number of repetitive responses, and in Experiment 2 this measure did not correlate with literacy. This result is consistent with a former study that reported no significant difference in number of repetitions between illiterate and (schooled) literate adults [17]. In fact, the only study that reported an influence of formal education on number of repetitive responses showed that participants with a higher level of formal education make more repetitions [24], probably because a higher level of education was also associated with increased total word production.

Thus, in terms of the several components that have been proposed to account for various experimental and individual differences effects in verbal fluency tasks [54,69,71], what our work show is that formal education and literacy impact on at least two of these components: in illiterate people or people with limited formal education, activation spreading from the cue to related items is more restricted, as the network, i.e., content knowledge, is itself more restricted, and the variety of retrieval cues is also more limited. But neither switching, nor monitoring the generated items, seems modulated by education and literacy. In addition, as do more formally educated people, adults with low levels of formal education, including illiterates, display a hierarchical organization of categories, with taxonomic clustering and retrieval by semantic subcategory.

Acknowledgments

Preparation of this article was supported by an Interuniversity Attraction Poles Grant (IAP 7/33) of the Belgian Science Policy Office (BELSPO, “Mechanisms of conscious and unconscious learning”). The first author is Research Director of the *Fonds de la Recherche Scientifique – FNRS*. At the time of data collection, the second and third Authors were PhD students at Universidade Federal de Santa Catarina, UFSC, Brazil.

References

- [1] Abwender DA, Swan JG, Bowerman JT, Connolly SW. Qualitative analysis of verbal fluency output: review and comparison of several scoring methods. *Assessment* 2001;8(3):323–36.
- [2] Arbuthnott K, Frank J. Executive control in set switching: residual switch cost and task-set inhibition. *Can J Exp Psychol* 2000;54(1):33–41.
- [3] Atran S, Medin DL. *The native mind and the cultural construction of nature*. Boston, MA: MIT Press; 2008.
- [4] Beck IL, McKeown MG. *Learning words well—a program to enhance vocabulary and comprehension*. *Read. Teach.* 1983;36:622–5.
- [5] Bledowski C, Kaiser J, Rahm B. Basic operations in working memory: contributions from functional imaging studies. *Behav Brain Res* 2010;214(2):172–9.
- [6] Blewitt P, Toppino TC. The development of taxonomic structure in lexical memory. *J. Exp Child Psychol* 1991;51(2):296–319.
- [7] Bousfield WA, Sedgewick CHW. An analysis of restricted associative responses. *J Gen Psychol* 1944;30(2):149–65.
- [8] Brucki SMD, Rocha MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Braz J Med Biol Res* 2004;37(12):1771–7.
- [9] Castro-Caldas A, Petersson KM, Reis A, Stone-Elander S, Ingvar M. The illiterate brain: learning to read and write during childhood influences the functional organization of the adult brain. *Brain* 2004;121(6):1053–63.
- [10] Charchat Fichman H, Santos Fernandes C, Nitirini R, Alves Lourenço R, Martins de Paiva Paradelo E, Carthery-Goulart MT, et al. Age and educational level effects on the performance of normal elderly on category verbal fluency tasks. *Dement Neuropsychol* 2009;3(1):49–54.
- [11] Collette F, Hogge M, Salmon E, Van der Linden M. Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience* 2006;139(1):209–21.
- [12] Downes JJ, Sharp HM, Costall BM, Sagar HJ, Howe J. Alternating fluency in Parkinson's disease: an evaluation of attentional control theory of cognitive impairment. *Brain* 1993;116(4):887–902.
- [13] Epker MO, Lacritz LH, Cullum CM. Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *J Clin Exp Neuropsychol* 1999;21(4):425–34.
- [14] Estes Z, Golonka S, Jones LL. Thematic thinking: the apprehension and consequences of thematic relations. In: Ross BH, editor. *Psychology of learning and motivation*. Burlington: Academic Press; 2011. p. 249–94.
- [15] Flores-Mendoza C, Widaman KF, Mansur-Alves M, da Silva Filho JH, Pasian SR, Furtado Schlottfeldt CGM. Considerations about IQ and human capital in Brazil. *Temas Psicol* 2012;20(1):133–54.
- [16] Friedman NP, Miyake A. The relations among inhibition and interference control functions: a latent-variable analysis. *J Exp Psychol: Gen* 2004;133(1):101–35.
- [17] Gonzalez da Silva G, Petersson KM, Faisca L, Ingvar M, Reis A. The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *J Clin Exp Neuropsychol* 2004;26(2):266–77.
- [18] Gruenewald PJ, Lockhead GR. The free recall of category example. *J Exp Psychol: Hum Learn Mem* 1980;6(3):225–40.
- [19] Gurd JM, Amunts K, Weiss PH, Zafiris O, Zilles K, Marshall JC, et al. Posterior parietal cortex is implicated in continuous switching between verbal fluency tasks: an fMRI study with clinical implications. *Brain* 2002;125(5):1024–38.
- [20] Henri JD, Phillips LH. Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Neuropsychol Cogn* 2006;13(3–4):529–51.
- [21] Kail R, Nippold MA. Unconstrained retrieval from semantic memory. *Child Dev* 1984;55(3):944–51.
- [22] Koren R, Kofman O, Berger A. Analysis of word clustering in verbal fluency of school-aged children. *Arch Clin Neuropsychol* 2005;20(8):1087–104.
- [23] Kosmidis MH, Tsapkini K, Folia V, Vlahou CH, Kiosseoglou G. Semantic and phonological processing in illiteracy. *J Int Neuropsychol Soc* 2004;10(6):818–27.
- [24] Kosmidis MH, Vlahou CH, Panagiotaki P, Kiosseoglou G. The verbal fluency task in the Greek population: normative data, and clustering and switching strategies. *J Int Neuropsychol Soc* 2004;10(2):164–72.
- [25] Kosmidis MH, Zafiri M, Politimou N. Literacy versus formal schooling: influence on working memory. *Arch Clin Neuropsychol* 2011;26(7):575–82.
- [26] Laine M, Niemi J. Word fluency production strategies of neurological patients: semantic and letter based clustering. *J Clin Exp Neuropsychol* 1988;10:28–35.

- [27] Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 2013;4:863.
- [28] Lin EL, Murphy GL. Thematic relations in adult's concepts. *J Exp Psychol: Gen* 2001;130(1):3–28.
- [29] Lombrozo T, Rehder B. Functions in biological kind classification. *Cogn Psychol* 2012;65(4):457–85.
- [30] Luria AR. *Cognitive development. Its cultural and social foundations*. Cambridge, MA: Harvard University Press; 1976.
- [31] Malt BC, Gennari S, Imai M, Ameal E, Saji N, Majid A. Where are the concepts? What words can and can't reveal In: Margolis E, Laurence S, editors. *The Conceptual Mind: New Directions in the Study of Concepts*, 2015, MIT. Cambridge, MA: MIT Press; 2014 (in press).
- [32] Mansfield AF. Semantic organization in the young child: evidence for the development of semantic feature systems. *J Exp Child Psychol* 1977;23(1):57–77.
- [33] Markman EM. The acquisition and hierarchical organization of categories by children. In: Sophian M, editor. *Origins of cognitive skills*. Hillsdale, NJ: Erlbaum; 1984.
- [34] Matute E, Montiel T, Pinto N, Rosselli M, Ardila A, Zarabozo D. Comparing cognitive performance in illiterate and literate children. *Int Rev Educ* 2012;58(1):109–27.
- [35] Mayr U. On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia* 2002;40(5):562–6.
- [36] Mayr U, Kliegl R. Complex semantic processing in old age: does it stay or does it go? *Psychol Aging* 2000;15(1):29–34.
- [37] Medin DL, Ross NO, Atran S, Cox D, Coley J, Proffitt JB, Blok S. Folkbiology of freshwater fish. *Cognition* 2006;99(3):237–73.
- [38] Mengarda EJ. Efeitos da alfabetização tardia no letramento, na linguagem e na cognição em adultos iletrados [The effects of late alphabetization on literacy acquisition, language and cognition in illiterate adults]. (PhD in Linguistics). Brazil: Universidade Federal de Santa Catarina; 2004.
- [39] Miyake A, Friedman NP, Emerson NJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: a latent variable analysis. *Cogn Psychol* 2000;41(1):49–100.
- [40] Morais J, Kolinsky R. Literacy effects on language and cognition. In: Bäckman L, von Hofsten C, editors. *Psychology at the turn of the millenium, vol. 1: cognitive, biological, and health perspectives*. Hove, UK: Psychology Press; 2002. p. 507–30.
- [41] Morais J, Cary L, Alegria J, Bertelson P. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 1979;7(4):323–31.
- [42] Newcombe F. *Missile wounds of the brain: a study of psychological deficits*. London: Oxford University Press; 1969.
- [43] Nisbett RE. *The geography of thought: how Asians and Westerners think differently... and why*. New York, NY: Free Press; 2003.
- [44] Petersson KM, Reis A, Askelöf S, Castro-Caldas A, Ingvar M. Language processing modulated by literacy: a network analysis of verbal repetition in literate and illiterate subjects. *J Cogn Neurosci* 2000;12(3):364–82.
- [45] Petersson KM, Reis A, Ingvar M. Cognitive processing in literate and illiterate subjects: a review of some recent behavioral and functional neuroimaging data. *Scand J Psychol* 2001;42(3):251–67.
- [46] Radeau M. Semantic priming between spoken words in adults and children. *Can J Psychol* 1983;37(4):547–56.
- [47] Raskin SH, Sliwinski M, Borod J. Clustering strategies on tasks of verbal fluency in Parkinson's disease. *Neuropsychologia* 1992;30(1):95–9.
- [48] Ratcliff G, Ganguli M, Chandra V, Sharma S, Belle S, Seaberg E, Pandav R. Effects of literacy and education on measures of word fluency. *Brain Lang* 1998;61(1):115–22.
- [49] Raven JC. *The Standard Progressive Matrices*. London, UK: Lewis; 1938.
- [50] Reis A, Castro-Caldas A. Illiteracy: a cause for biased cognitive development. *J Int Neuropsychol Soc* 1997;3(5):444–50.
- [51] Reis A, Guerreiro M, Petersson KM. Sociodemographic and neuropsychological characterization of an illiterate population. *Appl Neuropsychol* 2003;10(4):191–204.
- [52] Robert PH, Lafont V, Medecin I, Berthet L, Thaubay S, Baudu C, et al. Clustering and switching strategies in verbal fluency tasks: comparison between schizophrenics and healthy adults. *J Int Neuropsychol Soc* 1998;4(6):539–46.
- [53] Rosch E, Simpson C, Miller RS. Structural bases of typicality. *J Exp Psychol: Hum Percept Perform* 1976;2(4):491–502.
- [54] Rosen VM, Engle RW. The role of working memory capacity in retrieval. *J Exp Psychol: Gen* 1997;126(3):211–27.
- [55] Ross BH, Murphy GL. Food for thought: cross-classification and category organization in a complex real-world domain. *Cogn Psychol* 1999;38(4):495–553.
- [56] Rosselli M, Tappen R, Williams C, Salvatierra J, Zoller Y. Level of education and category fluency task among Spanish speaking elders: number of words, clustering, and switching strategies. *Aging Neuropsychol Cogn* 2009;16:721–44.
- [57] Saez-Zea C, Carnero-Pardo C, Gurpegui M. Nombres de personas: una prueba de fluidez verbal sin influencias socioeducativas [Names of persons: a verbal fluency test without socioeducational influences]. *Neurologia* 2008;23(6):356–60.
- [58] Santos GA, Lopes A, Neri AN. Escolaridade, raça e etnia: elementos de exclusão social de idosos [Education, race and ethnicity: elements of social exclusion among elders]. In: Neri AL, editor. *Idosos no Brasil: vivências, desafios e expectativas na terceira idade [Elders in Brazil: experiences, challenges and expectations in the third age]*. São Paulo: Fundação Perseu Abramo, SESC; 2007.
- [59] Scliar-Cabral L. Guia prático de alfabetização, baseado em princípios do sistema alfabético do Português do Brasil. [A practical guide to alphabetization, based on principles of the alphabetic system of Brazilian Portuguese]. São Paulo: Contexto; 2003.
- [60] Scribner S. Developmental aspects of categorized recall in a West African society. *Cogn Psychol* 1974;6(4):475–94.
- [61] Sharp D, Cole M, Lave C, Ginsburg HP, Brown AL, French LA. Education and cognitive development: the evidence from experimental research. *Monographs Soc Res Child Dev* 1979;44(1–2):1–112.
- [62] Shimura Barea K, Lessa Mansur L. Knowledge of semantic categories in normal aged: influence of education. *Dement Neuropsychol* 2007;1(2):166–72.
- [63] Strauss E, Sherman EMS, Spreen O. *A compendium of neuropsychological tests: administration, norms, and commentary*. Oxford: Oxford University Press; 2006.
- [64] Schwartz S, Baldo J, Graves RE, Brugger P. Pervasive influence of semantics in letter and category fluency: a multidimensional approach. *Brain Lang* 2003;87(3):400–11.
- [65] Tompkins CA, Bloise CGR, Timko ML, Baumgaertner A. Working memory and inference revision in brain-damaged and normally aging adults. *J Speech Hear Res* 1994;37(4):896–912.
- [66] Tröster A, Fields JA, Testa JA, Paul RH, Blanco CR, Hames KA, et al. Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia* 1998;36(4):295–304.
- [67] Troyer AK. Normative data for clustering and switching on verbal fluency tasks. *J Clin Exp Neuropsychol* 2000;22(3):370–8.
- [68] Troyer AK, Moscovitch M. Cognitive processes of verbal fluency tasks. In: Poreh AM, editor. *The quantified process approach to neuropsychological assessment*. New York, USA: Taylor & Francis; 2006. p. 142–62.
- [69] Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* 1997;11(1):138–46.
- [70] Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D. Clustering and switching on verbal fluency: the effects of focal frontal and temporal lobe lesions. *Neuropsychologia* 1998;36(6):499–504.
- [71] Unsworth N, Spillers GJ, Brewer GA. Variation in verbal fluency: a latent variable analysis of clustering, switching, and overall performance. *Q J Exp Psychol* 2011;64(3):447–66.
- [72] Ventura P, Kolinsky R, Querido JL, Fernandes S, Morais J. Is phonological encoding in naming influenced by literacy? *J Psycholinguist Res* 2007;36(5):341–60.
- [73] Verhaeghe A, Kolinsky R. O que os iletrados nos ensinam sobre os testes de inteligência. [What illiterates teach us about intelligence tests]. Lisbon: Fundação Calouste Gulbenkian- Fundação para a Ciência e a Tecnologia; 2006.
- [74] Williams RS, Patton Terry N, Metzger I. Kid Categories: a comparison of the category productions of LSES and MSES elementary school children. *Commun Disord Q* 2013;34(2):71–80.
- [75] Wixted JT, Rohrer D. Analysing the dynamics of free recall: an integrative review of the empirical literature. *Psychon Bull Rev* 1994;1(1):89–106.
- [76] Yassuda MS, Diniz BS, Flaks MK, Viola LF, Pereira FS, Nunes PV, et al. Neuropsychological profile of Brazilian older adults with heterogeneous educational backgrounds. *Arch Clin Neuropsychol* 2009;24(1):71–9.
- [77] Zec RF, Landreth ES, Fritz S, GAMES E, Hasara A, Fraizer W, et al. A comparison of phonemic, semantic, and alternating word fluency in Parkinson's disease. *Arch Clin Neuropsychol* 1999;14(3):255–64.