Section A. Tasks

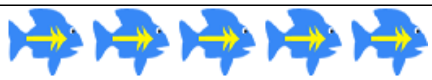
Inhibition

Response-Distractor Inhibition.

Flanker task. Initially introduced by Eriksen and Eriksen (1974), several previous studies have used the flanker paradigm to obtain measures of response-distractor inhibition with children and adolescents (Huizinga et al., 2006; Rueda et al., 2004; Simonds et al., 2007; Waszak et al., 2010). The present version of the task is an adaptation from Munro et al. (2006) and Rueda et al. (2004). The stimuli consisted of five cartoon fishes pointing to the right or the left, subtending a visual angle of 8º x 1.2º, horizontal and vertical, respectively. The color of the fishes was blue or pink, depending on whether participants attended to the central fish or the flanker fishes. A typical flanker task was administered in the blue condition, and the participants were to respond based on whether the central fish was pointing to the left or right, trying to ignore flanker fishes at the same time, by pressing the corresponding left or right key on the keyboard ('Z' and 'M,' respectively) with their index fingers of both hands. A reverse flanker task was performed in the pink condition, in which the participants responded to the flanker fishes' direction while ignoring the central fish.

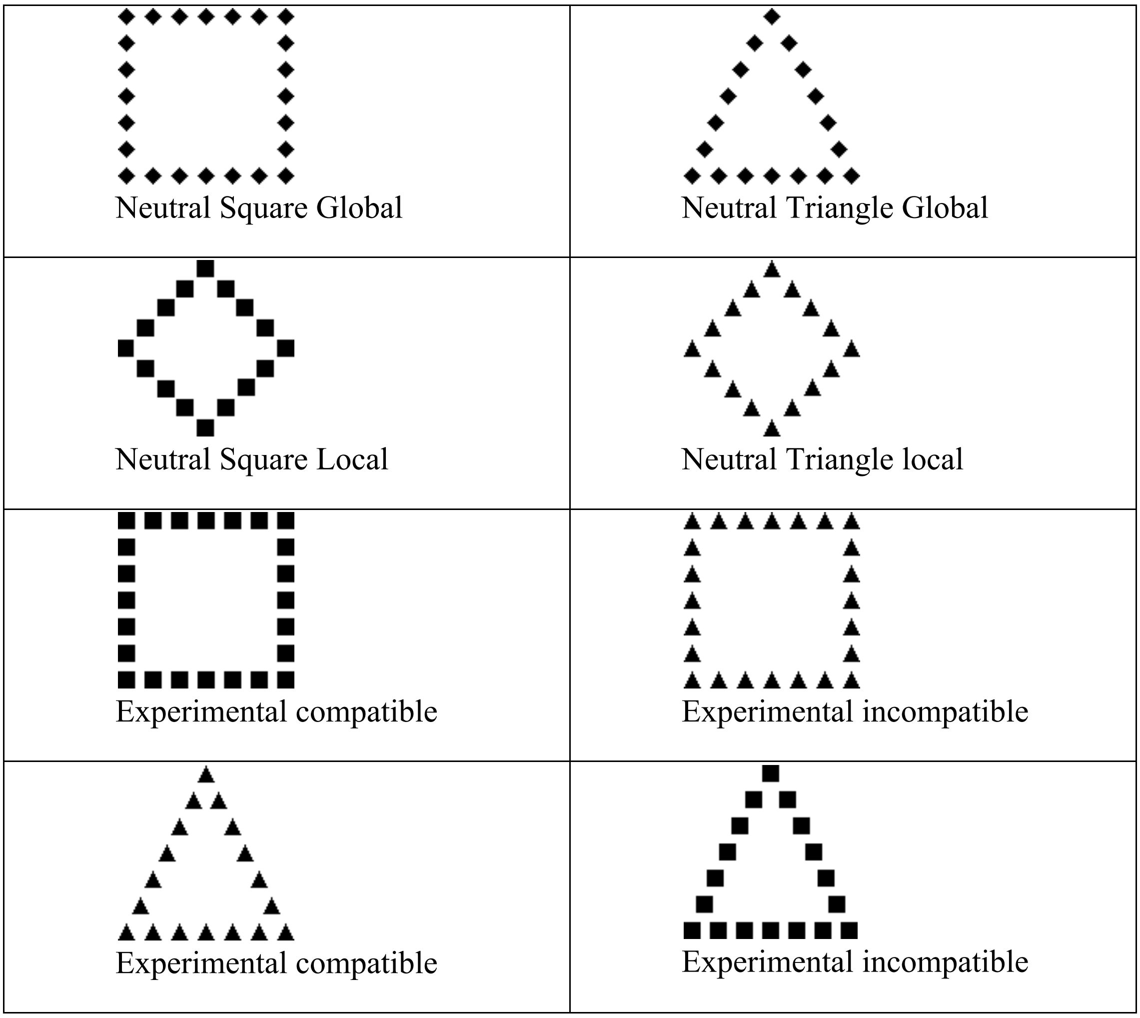
In three conditions, the target was flanked by three noise stimuli on each side: (a) on congruent trials, the flanking fishes were pointing in the same direction as the central one; (b) on incongruent trials, the flankers pointed in the opposite direction from the central fish; and (c) on neutral trials, the fishes to-be-ignored were replaced with geometrical figures without left-right defined direction (see Figure S1). Each trial started with a cross-shaped fixation point; 500 ms later, an array of five fishes appeared on the screen and remained until the participant responded. After the end of the trial, a pause of 500 ms ensued before the start of the next trial. The experiment started with a neutral blue block (32 trials; the first 8 were warm-up trials not included in the analysis), continued with a practice blue block in which participants had to respond to central fish (12 trials) and two experimental blue blocks (24 trials each). Then, the color of the fishes changed into pink, and the participants were instructed to respond to the flanker fishes instead of the central fish. Similarly to the blue condition, a neutral pink block (32 trials, first 8 were warm-up trials), a practice pink block (12 trials), and two experimental pink blocks (24 each) were successively administered. Finally, both central (blue) and flanker (pink) conditions were combined in the same blocks, with the fishes' color cueing each trial's target stimuli. A practice block (24 trials) and two experimental blocks (48 trials each) were applied under this alternating condition. Auditory and visual feedback was provided in a cartoon fashion to sustain a high attentional level (see Figure S1). Participants were instructed to respond as quickly as possible while making as few errors as possible. The dependent variable was the mean RT/percentage of hits in the flanker block's incongruent condition. The reliability estimates for RT and accuracy were .92 and .76, respectively, for the 2nd grade sample and .96 and .77 for the adolescent sample. The average duration was 20 minutes.

Figure S1.
Example of experimental and neutral conditions in the Flanker task.

	Experimental Center Congruent
	Experimental Center Incongruent
	Experimental Flanker Congruent
	Experimental Flanker Incongruent
	Neutral for Center condition
	Neutral for Flanker condition

Local-global Task. Initially developed by Navon, (1977), two previous developmental studies (Huizinga et al., 2006; Mondloch et al., 2003) administered a local-global paradigm to children and adolescents to measure response-distractor inhibition. The present version of the task was adapted from (Mondloch et al., 2003; Montoro et al., 2011). The stimuli consisted of large squares and triangles (or diamonds, in the neutral local condition) made up of small squares and triangles (or diamonds, in the neutral global condition). The global stimuli subtended a visual angle of 2.7º x 2.7º. The global squares were made up of 24 local elements, the global triangles were composed of 18 local elements, and the global diamonds consisted of 16 local elements. The local elements subtended a visual angle of 0.3º x 0.3º and consisted of solid black figures on white background. Eight different stimuli were presented in this experiment. The experimental procedure consisted of three phases. In the first phase, participants responded to the global figure; in the second phase, they responded to the local elements; and, in the third phase, an auditory cue previous to the stimuli (the spoken words' BIG' or 'SMALL') indicated to which level (global or local) the participants should respond in each trial. Participants were instructed to make their responses as quickly as possible while making as few errors as possible. Each trial started with a cross-shaped fixation; 500 ms later, a visual stimulus was presented at the center of the screen until responses. The participants had to indicate the identity of the stimulus at the level to be attended (global or local) by pressing a keyboard button ('Z' for the triangle; 'M' for square) using their index fingers of both hands (see Figure S2). At the bottom of the screen, a small triangle and a small square were displayed to remind this key assignment. Visual feedback was provided for every trial. The global phase started with a neutral block (32 trials; the first 8 were warm-up trials) that displayed large squares and triangles made up of smaller diamonds. After that, a global block practice (12 trials) and two experimental global blocks (24 trials) were administered. Similarly, the local phase started with a neutral block (32 trials; the first 8 were warm-up trials) that displayed large diamonds made up of smaller triangles or squares. Next, a local block practice (12 trials) and two experimental local blocks (24 trials) were administered. In the third phase, the participants performed an alternating global-local task that consisted of a practice block (24 trials) and two experimental blocks (48 trials). The dependent variable was RT/percentage of hits for incompatible conditions in local blocks. The reliability estimates for RT and accuracy were .90 and .84, respectively, for the 2nd grade sample and .96 and .81, respectively, for the adolescent sample. The average task duration was 20 minutes.

Figure S2.
Example of stimuli in Local-Global Task

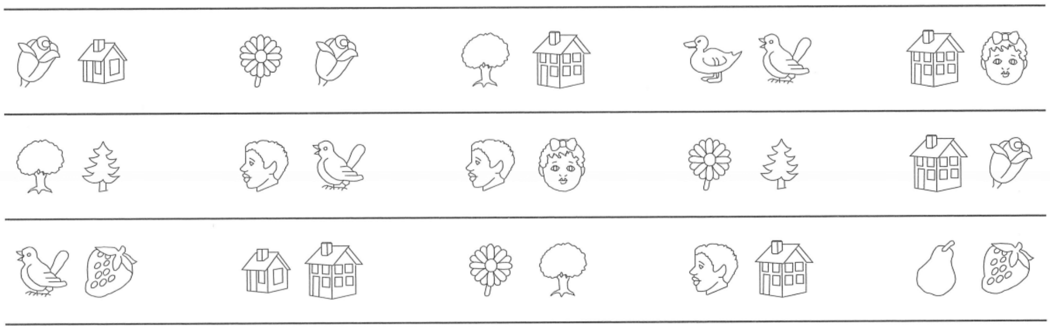


Receptive attention task. It is a scale of Cognitive Assessment System (CAS) (Naglieri & Das, 1997). This timed pencil-and-paper test measures the ability of selective attention requiring children to focus on relevant stimuli while ignoring irrelevant ones. 7-year-olds children were presented to 200 pairs of drawings per condition in four separate sheets comprising 50 pairs each. In the first condition, children needed to underline the physically identical drawings (e.g., they have to underline the drawing , but not  the drawing). For the second condition, they had to underline the drawings that belong to the same lexical category (that have the same name) (e.g., they have to underline the drawing  , but not the  drawing) (see Figure S3. Panel A). For 11 and 15-year-olds, a total of 400 pairs of letters per condition –physical identity and lexical identity– were presented in two separate sheets with 200 pairs each. For the first condition, participants needed to underline the physically identical pairs of letters (e.g., they had to underline *AA*, *aa,* but not *Aa*). For the second condition, participants had to underline the lexically similar pairs (e.g., they had to underline *Aa, AA, aa, but not Tr or tB*) (see Figure S3, panel B). The task was individually handed out following manual administration, also providing scoring rules for each group of age. The dependent variable was the number of correct answers minus the number of mistakes and the time to complete the test, a *base score* that accounts for both correctness and speed. Test-retest reliability reported in the manual is .89. The task duration was 8 minutes.

Figure S3.

Examples of Naglieri and Das stimuli for 7-year-olds (Panel A) and 11 and 15-year-olds (Panel B).

Panel A



Panel B

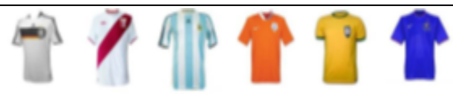
tB	**TT**	**Rb**	**ee**	**bN**
RR	**nE**	**nn**	**Tr**	**bt**
EA	**BB**	**TR**	**nb**	**aa**

Go-no go task. Three previous studies have used Go-no go tasks with samples of children and adolescents to measure response-distractor inhibition (Christ et al., 2006; Durston et al., 2002; Johnstone et al., 2005). The current version of the paradigm is an adaptation from (Christ et al., 2006). Two experimental conditions were provided: go and no-go. The no-go stimulus was the red t-shirt of the Spanish national football team, and the go stimuli were six t-shirts of other national football teams (i.e., Germany, Argentina, Brazil, France, Netherlands, and Peru) subtending approximately 4.3° horizontally and 5.3° vertically. On each trial, one of the t-shirts was centrally displayed. Participants were asked to press the space bar as quickly as possible when any stimulus except for the Spanish t-shirt appeared; in that case, participants had to avoid responding (see Figure

S4). After an interval of 1,000 ms, a new trial was presented. If a participant responded less than 100 ms after the presentation of a target (an anticipatory error), a visual message ("too fast, you cannot see the t-shirt yet") was displayed on the screen. In contrast, if a participant failed to respond within 1,500 ms (an inattentive error), a different visual message ("too slow, respond faster") was shown. If a participant responded on a no-go trial (a false alarm error), another visual message ("no response needed when you see the Spanish t-shirt") appeared. Following 49 go trials (neutral phase), six experimental blocks consisting of 40 trials (30 go and ten no-go trials) were administered. No-go stimuli was randomly included in 25% of the trials. At the end of each block, a break was offered. The dependent variable was the proportion of errors in no-go trials. The reliability estimates for accuracy and RT were .85 and .76, respectively, for the 2nd grade sample and .85 and .91, respectively, for the adolescent sample. Task's duration was about 15 minutes.

Figure S4.

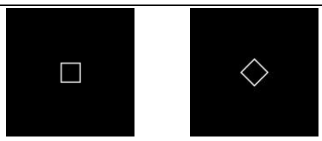
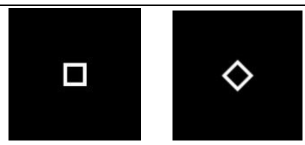
An example of stimuli in the Go-no go task.

No go	Go
	

Stop-signal Task. Initially designed by (Logan & Cowan, 1984) three previous developmental studies have included stop-signal tasks as an index of response-distractor inhibition (Huizinga et al., 2006; St Clair-Thompson & Gathercole, 2006; Van Den Wildenberg & Van Der Molen, 2004). We used the software called STOP-ITa, developed by (Verbruggen et al., 2008), to run the stop-signal paradigm. The task consisted of no-signal trials (75% of the trials), in which participants had to discriminate between a square and a diamond. Participants were instructed to respond as fast and accurately as possible. The rest of the trials (25%) were stop-signal trials in which the lines of the shapes would become thicker after a variable delay, and subjects were instructed to withhold their responses (see Figure S5). Each trial started with the presentation of the fixation sign, which is replaced by circles or diamonds after 250 ms Participants had to press the "Z" key for *square* and "M" for *diamonds* shapes. The stimulus remained on the screen for 1,250 ms. The interstimulus interval was 2,000 ms. On stop-signal trials, a stop signal was presented after a variable SSD (stop-signal delay). SSD was initially set at 250 ms and was adjusted continuously with the staircase tracking procedure: when inhibition was successful, SSD increased by 50 ms; when inhibition was unsuccessful, SSD decreased by 50 ms. The experiment consisted of two phases: a practice phase of 32 trials and an experimental phase of three blocks of 64 trials. Between blocks, subjects needed to wait for 10 s before they could start the next block. The dependent variable was the stop-signal reaction time (SSRT). Reported reliability from simulation studies over a sample of 100,000 participants and 375,000,000 trials was r =.53 (Verbruggen et al., 2019). The task's duration was about 15 minutes.

Figure S5.

Example of stimulus in Stop-signal Task

No signal	Stop signal
	

Stroop task. Designed initially by Stroop (1935), at least four previous works with children and adolescents have made use of the Stroop effect as a measure of response-distractor inhibition (Huizinga et al., 2006; Leon-Carrion et al., 2004; Rueda et al., 2005; Tamnes et al., 2010). We designed this version of the task following the recommendations provided by MacLeod (2006). Stimuli consisted of color items —asterisks (*****) and words "*red*" and "*blue*"— printed in blue and red colors in arial font and point size 32. Participants were then instructed to respond "red" or "blue" both vocally and by pressing "Z" and "M" keys with their left and right index finger as quick as possible. At the bottom of the screen, the letters "R" and "B" in gray color were displayed as reminders of this key assignment. First, a blank screen was presented for each trial lasting 500 ms. During this time, fixation point was shownfor the first 250 ms, . followed then by another blank for the last 250 ms. Next, the color item appeared in the center of the screen in the lower case (in the case of words). Finally, the color item was removed with the participant's response. The experiment consisted of three blocks of trials. The first block consisted of 60 neutral trials in which color items were five asterisks printed in blue or red colors. Participants were asked to state aloud and press the key associated with the color in which the asterisks were printed. Neutral trials were followed by 16 practice trials and then by 48 experimental trials. Half of the trials were congruent color items—that is, stimuli in which the color name matched the ink color (e.g., the word *blue* printed in blue color)—and the other half were incongruent trials in which the color name did not match the ink color (e.g., the word *blue* printed in red color). Congruent and incongruent conditions were intermingled randomly. The dependent variable was RT/percentage of hits in incongruent conditions. The reliability estimates for RT and accuracy were .93 and .78, respectively, for the 2nd grade sample and .96 and .67, respectively, for the adolescent sample.

Cognitive inhibition.

Updating information in Working Memory Task. We used an updating task originally devised by De Beni & Palladino (2004; Palladino et al., 2001), adapted for children and adolescents. The task provides some indexes of intrusions (previous-lists and same-list intrusions) that have been considered indexes of cognitive inhibition in previous developmental studies. Same-list intrusions have been considered an index of suppression of information in WM, and previous lists intrusions an index that accounted for proactive interference (Carriedo et al., 2016; De Beni & Palladino, 2004; Lechuga et al., 2006; Palladino et al., 2001, among others). The task contained 24 lists (20 experimental and four practice lists) of 12 words each. Each list included words to be recalled (relevant words), words to be discarded (irrelevant words), and filler words. The number of each kind of words in each list varied depending on the experimental condition (see Table S1). Thus, the number of relevant words in each list varied between 3 (low memory load) and 5 (high memory load). The number of irrelevant words varied between two (in the low suppression condition) and five (in the high suppression condition). Finally, the number of abstract filler words varied between 2 and 7. Target words (relevant and irrelevant) were familiar concrete nouns referring to body parts, objects, or animals that can be classified by size. Filler words were abstract nouns. Word frequency, concreteness, and familiarity were controlled. The final 24 lists were distributed in 4 experimental conditions of 6 lists. One list of each experimental condition was considered a practice list. The total number of words to be recalled across all condition lists was 80 (practice lists were excluded), 25 in each high memory load condition, and 15 in each low memory load condition. The dependent variable was the proportion of previous lists intrusions. The reliability estimates were .87 for the 2nd grade sample and .86 for the adolescent sample. The average task duration was 20 minutes.

Table S1.

Composition of the lists as a function of the experimental conditions

Low load/low suppression (5 Lists)	Low load/high suppression (5 Lists)	High load/low suppression (5 Lists)	High load/high suppression (5 Lists)
3 Relevant items 2 Irrelevant items 7 Filler items	3 Relevant items 5 Irrelevant items 4 Filler items	5 Relevant items 2 Irrelevant items 5 Filler items	5 Relevant items 5 Irrelevant items 2 Filler items

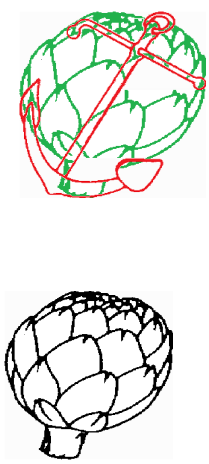
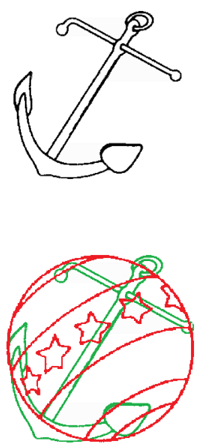
Negative Priming. The current version of the task is adapted from the previous work of Tipper (1985) and Pritchard & Neumann (2004). These previous studies used the negative priming effect as an index of cognitive inhibition in children (see also Rueda et al., 2005)[1]. The stimuli included line drawings from the Snodgrass & Vanderwart (1980)'s corpus (see Figure S6) representing objects or animals. These displays were organized into pairs consisting of a prime display followed by a probe display. An entirely overlapping red shape (non-target) and green shape (target) in each prime and probe display was revealed at the center of the screen. A black comparison shape appeared randomly, either above or under the overlapping shapes. The area occupied by the overlapping shapes subtended a visual angle of 2.4º x 2.4º, as well as the area occupied by the comparison shape. The distance between the outer edge of the

[1] Although there is considerable debate regarding whether negative priming is a measure of inhibition or of episodic retrieval (Mayr & Buchner, 2007), the consideration that the negative priming effect is accounted for by inhibitory processes is a widely accepted interpretation (see Rueda et al., 2005). We selected negative priming as a measure of cognitive inhibition based on two different criteria: a) previous research results with children (Pritchard & Neumann, 2004); and b) the empirical correlation found in our sample in pilot studies between negative priming and other inhibition-related measures (.45 with previous list intrusions in updating working memory tasks, .67 with flanker task; .59 with local-global task and - .49 with receptive attention task; all *ps* <.01).

overlapping and comparison shapes was about 0.4º. The red shape was described as a distractor, and participants were advised to ignore it to their best. Participants were then instructed to decide if the green target shape matched the black comparison shape by pressing "Z" and "M" keys with their left and right index fingers for "different" and "same" responses, respectively, as quick as possible. At the bottom of the screen, the words "different" and "same" were displayed as reminders of this key assignment. Half of the experiment trials were control trials (in which the red prime distractor and the black target probe were different items), and half were ignored repetition trials (the red prime distractor and the black target probe were the same items). Each session began with a set of 20 practice trials. Then, information about mean reaction times and hit rate during practice block was displayed on the screen. The experimenter decided if the participant needed to repeat the practice block or was ready to start the experimental phase as a result of his/her performance. The experimental phase consisted of 216 trials, divided into four blocks of 54 each. Each experimental trial consisted of a couplet of prime-probe displays. It began with a 2000 ms presentation of a cross-shape fixation point followed by a blank screen for 500 ms before the prime display. In the prime display, the overlapping red and green shapes and the black comparison shape appeared simultaneously. Participants had to decide which green shape (to be attended) was the same as the black shape pressing the keys "same" or "different." The prime display remained on the screen until response, followed by a 500 ms interval in which a blank screen was displayed before the probe display. The overlapping red and green shapes and the black comparison shape were presented simultaneously and remained until a response in the probe display. In the ignore repetition condition, the stimulus to be ignored in the prime display (red) was the stimulus to be attended in the probe display (green). The stimulus to be ignored in the prime trial (red) did not become the stimulus to be attended (green) in the probe trial in the control condition. Feedback was provided for prime and probe responses only during the practice block. The dependent variable was RT/percentage of hits for the ignore repetition trials in the probe condition. The reliability estimates for RT and accuracy were .96 and .88, respectively, for the 2nd grade sample and .99 and .86, respectively, for the adolescent sample. The average task duration was 20 minutes.

Figure S6.

An example of an ignored repetition prime–probe pair requiring a "same" response both in prime and probe displays.

Prime	Probe
	

Note. The comparison shapes on the bottom and the top of each display were black, the target shapes in the center were green, and the distractor shapes in the center were red. Children were asked to respond "same" if the green and black items matched and "different" if they did not match in both displays.

Processing Speed.

We used the RT/percentage of hits in the neutral condition of the Stroop task. This measure has been previously used in a study on the structure and organization of executive functioning across development by (Huizinga et al., 2006).

Working Memory Capacity

Reading Span Test for children (PAL-N). The current version of the task is an adaptation for children of Daneman & Carpenter's (1980) Reading Span Task (Carriedo & Rucián, 2009). It measures the ability to process and store verbal information simultaneously, with no option to review strategies. Participants had to read aloud each sentence at its own pace and remember the last word of the sentence. The task consisted of 48 phrases (6 training and 42 experimental) grouped into levels of 2, 3, 4, and 5 sentences, with three series of sentences for level. At the end of each series, participants had to remember the last word of each sentence in the same order as presented. Sentences were presented individually in the center of a computer screen only for the time needed to read the sentence. As soon as the participant finished reading, a new sentence appeared. When the series finished, a question mark appeared on the screen, asking the participant to remember the words. The task ended when the participant was unable to remember any of the series of a given level. The dependent variable was reading span, considered at the level at which the participant had correctly answered at least 2 of the three series. The reliability and validity have been shown through hundreds of studies, ranging from .70–.90 for span scores (see Conway et al., 2005 for a review). The average task duration was 15 minutes.

Table S2.

Example of the three series of sentences of level 2. In bold, the word to be remembered in the same order as presented.

Series 1

Harry Potter is the most famous character in **cinema**

The summer was so cold that many people changed their **plans**

Series 2

Yesterday, the whole town listened to the mayor's **speech**

When flying, the pigeon lost a **feather**

Series 3

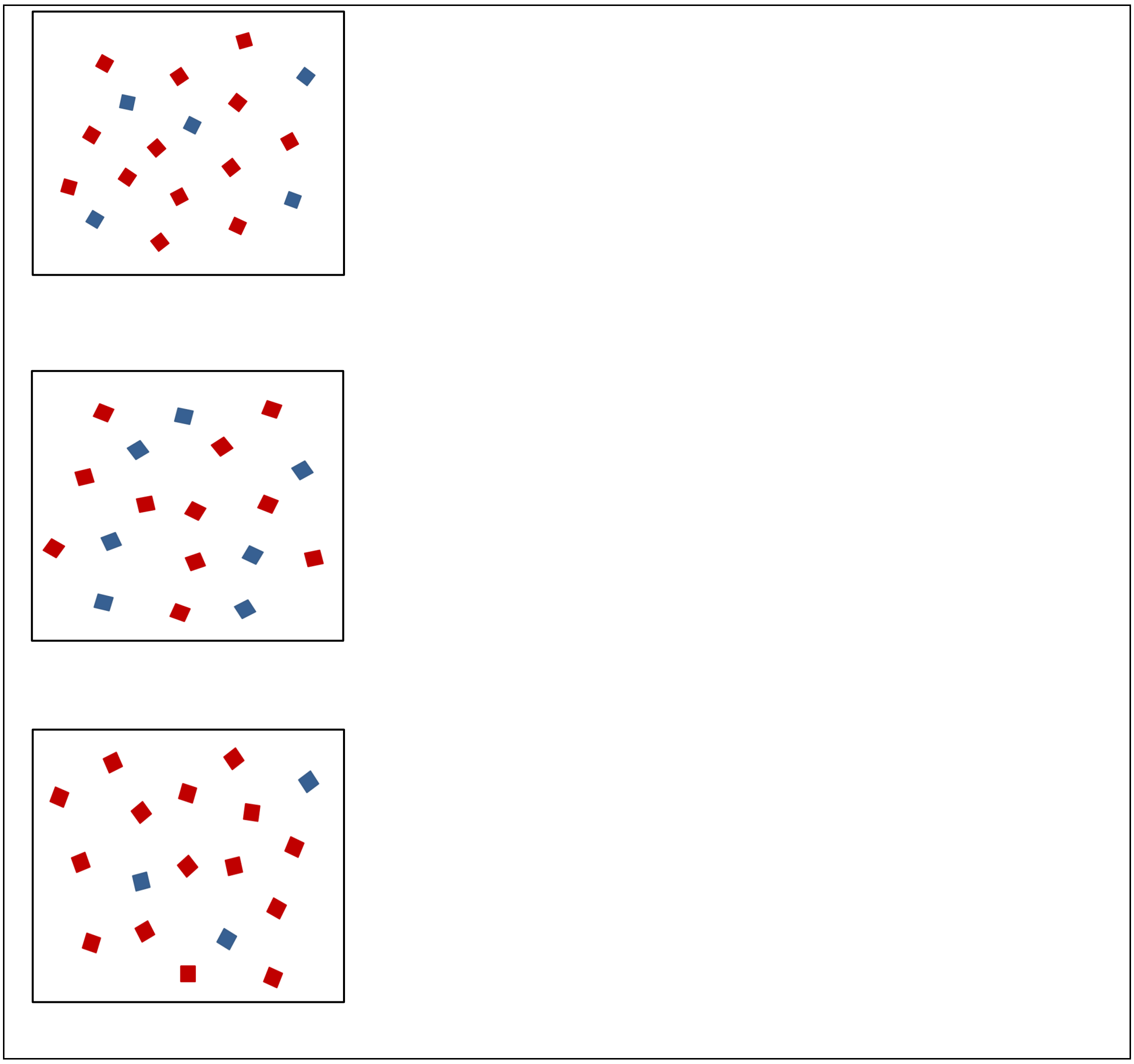
With his beautiful eyes he directed a deep look at his **mother.**

When we saw that he had a fever we went to tell the **doctor**

Counting Span Task (CST) (Case et al., 1982). This test measures the ability to process and store nonverbal information simultaneously, with no option to use review strategies. It has the same structure as Daneman and Carpenter's Reading Span Test with the difference that participants had to process visual information (counting and pointing out geometric figures) rather than verbal information. It consists of 48 visual displays (6 training and 42 experimental) grouped into levels. Each level contains 3 series of 2, 3, 4, and 5 displays each. Each display is formed by 18 squares, of which a minimum of 3 and a maximum of 8 are blue, and the rest are red squares to act as distractors (see Figure S7). Participants had to count the blue squares of each display without a pause. At the end of the series, they had to remember the number of squares counted in the same order in which they were presented. CST, in its original format, was handed out manually. However, we presented the stimuli on a computer screen. Initially, 42 experimental slides and six practice slides were built with 18 red and blue squares on a white background. All squares were randomly placed on each slide regarding location and rotation angle (greater or lesser rotation). The order of presentation of the slides was randomized so that patterns of numbers that facilitated remembrance did not appear in the same series (e.g., 4-5-6; 3-5-7). Each configuration of blue and red squares was presented individually on a computer screen for as long as it took to the participant to count them. When the participant finished counting, a new display appeared on the screen. To prevent repetition strategy, children had to point and count aloud each of the blue squares. A question mark appeared on the screen when the series finished, asking the participant to remember the numbers. The task ended when the participant was unable to remember any of the series of a given level. The reliability and validity have been shown through hundreds of studies, ranging from .70–.90 for span scores (see Conway et al., 2005 for a review). The average task duration was 10 minutes.

Figure S7.

Example of one series of displays of level 3. Participants had to point out and counting blue squares. Thus, they must recall “5”, “7”, “3”



Updating information in Working Memory Task. (See depiction in the inhibition tasks epigraph). For updating, the dependent variable was the proportion of correct responses in immediate recall (index of updating). The reliability estimates were .87 for the 2nd grade sample and .86 for the adolescent sample.

Cognitive flexibility

Flanker. Alternating block where both central (blue) and flanker (pink) conditions were combined, with the fishes' color cueing each trial's target stimuli. A practice block (24 trials) and two experimental blocks (48 trials each) were applied under this alternating condition. Auditory and visual feedback was provided in a cartoon fashion to sustain a high attentional level (see Figure S1). The dependent variable was the mean RT/percentage of hits in change condition.

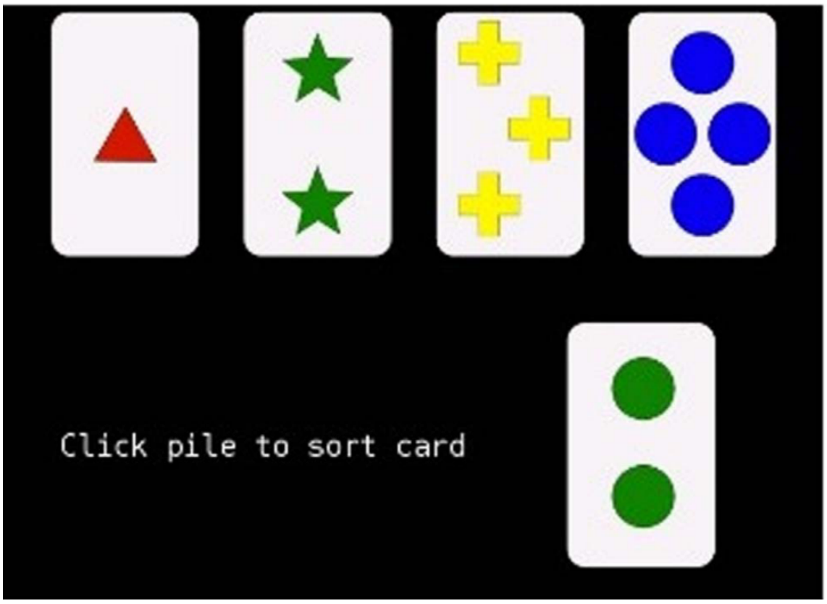
Local-Global. Alternating block where local and global conditions were combined. Auditory cues cueing each trial's target stimuli. It consisted of a practice block (24 trials) and two experimental blocks (48 trials). The dependent variable was RT/percentage of hits for the change conditions in the alternating block.

Wisconsin Card Sorting Test. We used a version of the classical Wisconsin task based on Berg (1948) with 64 trials, pBCST-64 (Mueller, 2011). After the presentation of the instructions, the first trial began with a screen on which a row of four piles cards across the top. Each pile has a different color, shape or number and a another card will appeared at the bottom right (see Figure S8). The participant has to classified the series of 64 cards appearing at the bottom. Participants had to choose the correct response by clicking with the mouse one of the four standard cards. The correct response depends upon a rule that the participants don't know and that they have to infer based on the feedback. Immediately after the response, a message of "Right!" or "Wrong!" appeared on the screen for 500 ms and, from this feedback. The rule changed without warning after ten consecutive correct trials. The total number of trials was 64. The experimental procedure was controlled by computer and the participants performed the task individually. The strong correlations found in Fox et al. (2013) between performance on the 128-trial and 64-trial versions of the BCST for total errors, perseverative responses, and perseverative errors suggest that the shorter 64-trial BCST is an acceptable alternative to the longer, full-length test.

We analized RTs of every trial following the same procedure that Somsen et al., (2000). We computed the average of the 5 trials just before the change of rule and the average of the five trials inmediately after the change of rule. We used two different indicators for this task, the classic perseverative errors measure and the mean RT time of the trials just after the change of rule. The reliability estimates were .90 for the 2nd grade sample and .92 for the adolescent sample. The average time to complete the task was 12 minutes.

Figure S8.

Example of display for the WCST



Long-term memory retrieval

Updating information in Working Memory Task. (See depiction in the inhibition tasks epigraph). After finishing updating task, in a second phase, participants had to perform a delayed recall task in which they are asked to recall as many words as they can from the presented in the first phase. To avoid a recency effect, participants had to perform an interpolating task (a backward counting task) between the updating and the delayed recall tasks. The dependent variable was the proportion of total words, correct and incorrect words recalled

Fluid intelligence

Raven Standard Progressive Matrixes (SPM, General) (Raven et al., 1998). The general scale has 60 items organized in five sets (A to E) with 12 items in each set. In each item, children had to complete a series of complex spatial figures using analogical reasoning. The final score was the sum of the number of correctly solved problems. The reliability estimates were .95 and .93, respectively, for the 2nd grade and adolescent samples.

Processing speed

Stroop task. We used the RT/percentage of hits in the neutral conditions. The reliability estimates for RT and accuracy were .93 and .78, respectively, for the 2nd grade sample and .96 and .67, respectively, for the adolescent sample.

Long-term memory retrieval

Updating information in Working Memory Task. (See depiction in the inhibition tasks epigraph). For long-term memory retrieval, the dependent variable was the proportion of correct responses in delayed recall.

Section B. Results

Table S3

Categories of animals

Wild animals: aardvark, alligator, anaconda, anteater, antelope, ape, armadillo, auk, baboon, badger, bat, bear, beaver, Bengal tiger, bison, boa, bobcat, brown bear, buffalo, camel, capybara, caribou, chameleon, cheetah, chimpanzee, cobra, cougar, crocodile, deer, dormouse, elephant, elk, emu, fallow deer, fawn, fox, gazelle, giraffe, gorilla, grizzly bear, hedgehog, hippopotamus, hyena, impala, jackal, jaguar, kangaroo, koala, lemur, leopard, lion, lioness, lynx, mammoth, marsupial, meerkat, mongoose, monkey, moose, mouflon, mountain lion, musk ox, opossum, orangutan, ostrich, panda bear, pangolin, panther, platypus, polar bear, porcupine, primate, puffin, puma, python, raccoon, reindeer, rhinoceros, roe deer, sable, shrew, skunk, sloth, snake, squirrel, tiger, viper, white tiger, wild boar, wildebeest, wolf, zebra
European wild animals: badger, bear, brown bear, buffalo, deer, fawn, fox, hedgehog, lynx, monkey, mouflon, porcupine, wild boar, wildcat, wolf
Non-European wild animals: ape, armadillo, baboon, Bengal tiger, bison, buffalo, camel, capybara, chimpanzee, cougar, elephant, giraffe, gorilla, hippopotamus, hyena, jaguar, kangaroo, koala, leopard, lion, monkey, orangutan, panda, pangolin, panther, platypus, primate, raccoon, skunk, tiger, zebra, white tiger
American wild animals: alligator, badger, beaver, caribou, cougar, grizzly bear, jaguar, moose, mountain lion, panther, raccoon, skunk, squirrel
African wild animals: antelope, camel, zebra, aardvark, ape, bee, cheetah, chimpanzee, crocodile, elephant, gazelle, giraffe, gorilla, hippopotamus, hyena, impala, jackal, lemur, leopard, lion, meerkat, mongoose, monkey, rhinoceros, tiger, wildebeest
Australian wild animals: emu, kangaroo, koala, platypus, possu
Arctic wild animals: auk, bear, caribou, musk ox, penguin, polar bear, puffin, reindeer, sea lion, seal, walrus
Asian wild animals: Bengal tiger, elephant, monkey, orangutan, panda bear, pangolin, panther, silkworm, skunk, white tiger
Farm animals: baby bird, bee, bird, bull, calf, canary, cat, chick, chicken, chihuahua, cow, Dalmatian, dog, donkey, duck, ferret, fish, foal, frog, German shepherd, goat, goose, greyhound, hamster, hen, horse, hound, lamb, llama, mare, marmot, mastiff, mouse, mule, ostrich, ox, parakeet, pheasant, pig, pigeon, piglet, pony, quail, rabbit, ram, rat, rooster, sheep, sheepdog, snail, swallow, tadpole, turkey, worm
Water animals: alligator, anchovy, auk, barbel, barracuda, beaver, blue shark, blue surgeonfish, blue whale, bogue, bonito, butterfly fish, carp, catfish, clam, clown fish, conch, crab, crocodile, cuttlefish, death watch, dolphin, duck, duckling, eel, emperor, fish, fish frog, flamingo, frog, grouper, hake, hammerhead, jellyfish, leech, lobster, manatee, mussel, Norway lobster, octopus, orca, otter, oyster, penguin, pike, piranha, plankton, platypus, prawn, puffer fish, puffin, ray fish, roosterfish, salmon, sardine, sea bass, sea bream, sea cucumber, sea lion, sea urchin, seahorse, seal, shark, shellfish, shrimp, snail, snake, sole, sperm whale, spider fish, sponge, squid, star, starfish, swan, swordfish, tadpole, toad, trout, tuna, turtle, walrus, water snake, whale, white-blue whale
Mountain/forest animals: baboon, badger, bat, bear, bee, bird, buck, chamois, condor, crow, deer, eagle, earthworm, elk, fallow deer, fawn, ferret, fox, frog, goat, golden eagle, gorilla, grasshopper, hare, hawk, hedgehog, ibex, jaguar, lizard, marsupial, mole, mouflon, otter, owl, partridge, pheasant, puma, quail, rabbit, reindeer, roe deer, skunk, sloth, slug, snail, snake, spider, squirrel, turtle, vulture, wild boar, wolf, wolverine, worm
Jungle/savanna animals: anaconda, ant, ape, baboon, bee, boa, chameleon, chimp, elephant, giraffe, hippopotamus, insect, jaguar, lion, mosquito, orangutan, ostrich, panther, python, rat, rhinoceros, snake, spider, tarantula, toucan
Birds: agaporni, bearded vulture, bird, blackbird, bunting, canary, chaffinch, chickadee, chicken, cockatoo, condor, crane, crow, diamond, duck, eagle, egret, emu, finch, flamingo, golden eagle, goldfinch, goose, goshawk, greenfinch, gull, harrier, hawk, hen, hummingbird, jay, kea, kestrel,

kingfisher, kiwi, lark, macaw, magpie, nightingale, nymph, ostrich, owe, owl, parakeet, parrot, partridge, peacock, pelican, penguin, pheasant, pigeon, pingan, puffin, quail, rhea, robin, rooster, sparrow, stork, swallow, swan, thrush, toucan, turkey, turtledove, vulture, woodpecker
Flying animals: agaporni, ant, bat, bearded vulture, bee, beetle, bird, blackbird, bumblebee, bunting, butterfly, canary, chaffinch, chick, chickadee, cicada, cockatoo, condor, crane, cricket, diamondback, dragonfly, duck, eagle, egret, emu, firefly, flamingo, fly, gnat, golden eagle, goldfinch, goose, goshawk, grasshopper, greenfinch, gull, harrier, hawk, hen, horsefly, hummingbird, insect, jay, kea, kestrel, kingfisher, kiwi, ladybug, lark, locust, macaw, magpie, mantis, mosquito, nightingale, nymph, owe, owl, parakeet, parrot, partridge, peacock, pelican, pheasant, pigeon, pingan, praying mantis, puffin, quail, raven, rhea, robin, rooster, sapsucker, sparrow, stick insect, stork, swallow, swan, termite, thrush, totovia, toucan, turkey, turtle dove, vulture, wasp, woodpecker
Bovines: antelope, bison, buck, buffalo, bull, calf, chamois, cow, kudu, mouflon, musk ox, oryx, ox, roe deer, wildebeest, yak
Canines: bitch, chihuahua, coyote, Dalmatian, dog, fox, German shepherd, greyhound, hound, hyena, jackal, mastiff, shepherd dog, wolf, wolfhound
Felines: Bengal tiger, cat, cheetah, jaguar, leopard, lion, lynx, mountain lion, ocelot, panther, puma, tiger, white tiger
Cervids: caribou, deer, elk, fallow deer, fawn, gazelle, impala, pudu, red deer, reindeer, roe deer, wildebeest
Insects/arachnids/worms: ant, aphid, bedbug, bee, beetle, beetle cutter, bumblebee, butterfly, caterpillar, centipede, cicada, cobbler, cockroach, cricket, dragonfly, earthworm, earwig, firefly, flea, fly, grasshopper, horsefly, insect, ladybug, larva, leech, locust, louse, mite, mosquito, moth, praying mantis, scorpion, silkworm, spider, stick insect, tarantula, termite, tick, wasp, woodworm, worm
Mollusks: clam, cuttlefish, mollusk, mussel, octopus, oyster, slug, snail, squid
Mustelids: badger, ferret, mink, otter, shrew, skunk, stoat, weasel
Insectivores: anteater, armadillo, bat, bird, blackbird, chameleon, dragonfly, frog, gecko, goldfinch, hedgehog, kiwi, ladybug, lizard, mantis, meerkat, mole, mongoose, newt, nightingale, salamander, scorpion, shrew, sloth, snake, sparrow, spider, starling, swallow, tarantula, thrush, toad, wasp, woodpecker
Primates/apes: ape, chimpanzee, gorilla, human, lemur, mandrill, marmoset, monkey, orangutan, primate
Leporidae: guinea pig, hare, rabbit
Reptiles/amphibians: alligator, anaconda, boa, chameleon, cobra, crocodile, dinosaur, freshwater turtle, frog, gecko, iguana, Komodo dragon, lizard, newt, python, reptiles, salamander, snake, toad, turtle, viper, water snake
Rodents: badger, beaver, capybara, chinchilla, dormouse, ferret, gerbil, guinea pig, hamster, marmot, marten, mole, mouse, muskrat, polecat, porcupine, rat, shrew, squirrel, weasel
Camelids: alpaca, camel, dromedary, guanaca, llama, vicuna
Ruminants: antelope, bison, buffalo, bull, caribou, cow, elk, gazelle, giraffe, goat, impala, kudu, mouflon, musk ox, okapi, oryx, ox, ram, reindeer, roe deer, sheep, wildebeest, yak
Bears: bear, brown bear, grizzly bear, polar bear, panda bear
Pack animals: ass, camel, donkey, dromedary, horse, llama, mare, mule, ox, ya
Animals used for fur: bear, beaver, chinchilla, cow, crocodile, fox, leopard, mink, otter, rabbit, sheep, snake
Pets: agapornis, anaconda, bird, bitch, blue surgeon, boa, butterfly fish, canary, cat, chameleon, Chihuahua, chinchilla, cobra, cockatoo, Dalmatian, dog, ferret, fish, gerbil, German shepherd, golden, goldfinch, greyhound, guinea pig, hamster, hedgehog, hound, iguana, little bird, lizard, mastiff, parakeet, parrot, pig, pigs, python, rabbit, retriever, sheepdog, silkworm, snake, tarantula, turtle, viper
Animals that are eaten: anchovy, barbel, bird, boar, bogue, bonito, bream, buffalo, bull, calf, camel, carp, chick, chicken, ciochinillo, clam, conch, cow, crab, cuttlefish, deer, dolphin, dromedary, duck, eel, fallow deer, fawn, fish, frog, goat, goose, grub, hen, horse, insect, lamb, little birds, lobster, mare, mouflon, mussel, Norway lobster, octopus, ox, oyster, partridge, pheasant, pig, pike, puffer fish,

quail, rabbit, ram, ray, roe deer, rooster, salmon, sardine, sea bass, sea bream, sea cucumber, shark, shellfish, shrimp, skate, snail, squid, trout, tuna, turkey, turtle, whale, wild boar
Animals and offspring: horse, mare, foal, goat, billy goat, goat, kid, hen, rooster, chicken, chick, pullet, sheep, lamb, ram, frog, tadpole, cow, calf, heifer, ox, bull, deer, fawn, pig, piglet, suckling pig, chick, goat, kid, lamb, chick, foal, calf, tadpole, pig, suckling pig, wolf, wolfling, chick, chick, she-wolf
Male and female: horse, mare, ram, sheep, lion, lioness, cow, bull, ox, goat, kid, goat, male goat, kid, wolf, she-wolf, monkey
Animals that begin/end with the same syllable: quail, partridge, rooster, hen, lion, lioness, fly, mosquito, puffer fish, clown fish, rat, mouse, badger, ferret, rhea, wildebeest, emu, lobster, crayfish, chameleon, bee, sheep, sloth bear, bear, python, pony, dove, quail, wolf, wolverine, fish, chicken, chick, chicklet, chick, wolf, bear, goat, goat, goat goat, mountain goat
Strong associative pairs: bee-bee-wasp, snail-babosa, crocodile-alligator, conch-shell-clam, rabbit-hare, snake-snake, shrimp-lobster, shrimp-prawn, lion-tiger, leopard-cheetah, scolopendra-snake, dog-cat, fly-mosquito, frog-toad, rat-mouse, cow-ox, shark-whale, whale-dolphin, mammoth-elephant, mouse-hamster, mouse-bat, mouse-bat, bonito-tuna, fox-cygnet, wild boar-pig, elephant-mouse, mouse-hamster, conch-mussel, wild boar-pig, snake-python, mouse-python, snake-snake, snail-snail, cat-mouse, sheep-bee, dog-hare
Prehistoric animals: dinosaur, mammoth
Mythological animals: cyclops, dragon, nymph, unicorn, mermaid

Data preparation and trimming procedure for timing executive tasks

Following Friedman et al. (2009) procedure RTs < 200 ms for the timing tasks were eliminated. Only trials on which correct responses were given were analyzed. RTs were analyzed using a trimming procedure robust to nonnormality (Wilcox & Keselman, 2003). Scores above three standard deviations (SDs) from the mean age group were replaced with values of the group mean plus three SDs: 6.6%, 5.3%, and 6.6% of the experimental scores for 7-, 11-, and 15-year-olds, respectively. We used the Inverse Efficiency Index (IEI), in which the RTs are divided by the percentage of hits (Townsend & Ashby, 1983) to account for the speed-accuracy trade-off. Additionally, to avoid mathematical problems with RTs' differential scores. Times to stop the response in the Stop signal task were not trimmed because these times do not depend on RT.

Section C. Clustering algorithm´s pseudo-code

The pseudocode of the clustering algorithm used in the automatic coding tool is the following:

```
------------------------------------------------------------
Algorithm 1: Clustering process
------------------------------------------------------------
Input: dataMatrix (3xN matrix with columns subject, property, order), t (threshold for clusters)
Output: clusters (a set with all final clusters)
1: cleanData = cleanningProcess(dataMatrix) #Cleaning process, eliminating repeated properties

######### Similarity process ######
2: numSubjects = number of subjects in the data
3: numProperties = Number of unique words mentioned listed by subjects
4: simM = Square similarity matrix of size numProperties x numProperties
5: for i from 1 to numSubjects:
6:       N_i = Number of words for subject i
7:       for x from 1 to N_i:     #iterating over the words mentioned by i
8:            for y from x+1 to N_i:  #to calculate the similarity between words for indexes x and y
9:                 d_i(x,y) = y-x-1   #number of words between x and y
10:                Add for words corresponding to indexes x and y in simM: (N_i - d_i(x,y) -1) / (N_i - 1).
11: simM = simM / numSubjects #Average over the number of subjects

######### Clustering process ######
12: clusters = set of all final clusters #empty set at the beginning
13: while(assignProperties < numProperties): #If a property has not been assigned to a cluster
    repeat the cycle
14:      Let x and y the indexes of the most similar properties from simM
15:      if simM(x,y) ≥ t:
16:           if (x belong to clusters and y does not belong to clusters):
17:                y is assigned to the same cluster than x.
18:           if (y belong to clusters and x does not belong to clusters):
19:                x is assigned to the same cluster than y.
20:           if (x and y do not belong to clusters):
21:                Create new cluster with x and y and add it to clusters.
22:      else:
23:           if (x or y does not belong to clusters):
24:              Assign x or y to cluster infrequent.
25:      simM(x,y) = 0.0
26: return clusters
```

The algorithm receives the original data matrix consisting of 3 columns, where each row has the number of the subject, a listed word, and the order of the listed word. The algorithm starts by eliminating the repeated words at line 1. Then, from lines 2 to 11, it calculates the similarity matrix among the words. The main loop of this process starts at line 5 by iterating over each subject. Line 6 extracts the number of words for subject *i*. Lines 7 to 10 iterate over all combinations of the words mentioned by subject *i*. Inside this loop, line 9 calculates the similarity between two specific words, and then line 10 updates the similarity matrix. Finally, line 11 calculates the final similarity as the average among all subjects. Lines 12 to 25 apply the proposed clustering algorithm. The main loop from lines 13 to 25 is repeated until all words have been assigned to a cluster. The loop starts at line 14 by searching for the most similar pair of words *x* and *y*. Lines 15 to 21 assigns words *x* and *y* to a specific cluster if their similarity is greater than or equal to a threshold *t* and at least one of them does not belong to a cluster. Otherwise, if the words have not been assigned to any existing cluster, they are assigned to the infrequent cluster. To avoid an infinite loop, line 25 overwrites the similarity between the words *x* and *y* corresponding to the similarity with indices *x* and *y* in matrix *M* to 0.0. In this way, it is impossible to select again the same pair of words *x* and *y* as the most similar pair of words. Finally, line 26 returns the final set of clusters contained in the variable labeled clusters.

Note that the new clustering algorithm, included in the automatic coding tool, will not put in the same cluster elements that are generated after a shift. For example, let's consider a typical sequence observed in the data: "dog", "cat", "lion", "tiger". In this sequence, the first two words correspond to pets, but then a shift is made changing to a cluster related to felines. For this example, classical clustering algorithms will create a cluster with those 4 words. In contrast, our algorithm will be able to separate "dog" and "cat" in one cluster and "lion" and "tiger" in another cluster.

Section D. Determination of similarity threshold (*t*) for the clustering algorithm

To determine the threshold (*t*) used in the clustering algorithm for the *clothing items* and *foods* categories we used the same criterion employed for *animals*. We simply plotted the number of clusters and shifts versus the threshold for the three age brackets and adopted the value that maximized those variables. Figures S9 and S10 show the number of clusters and shifts vs. *t* for *clothing items*.

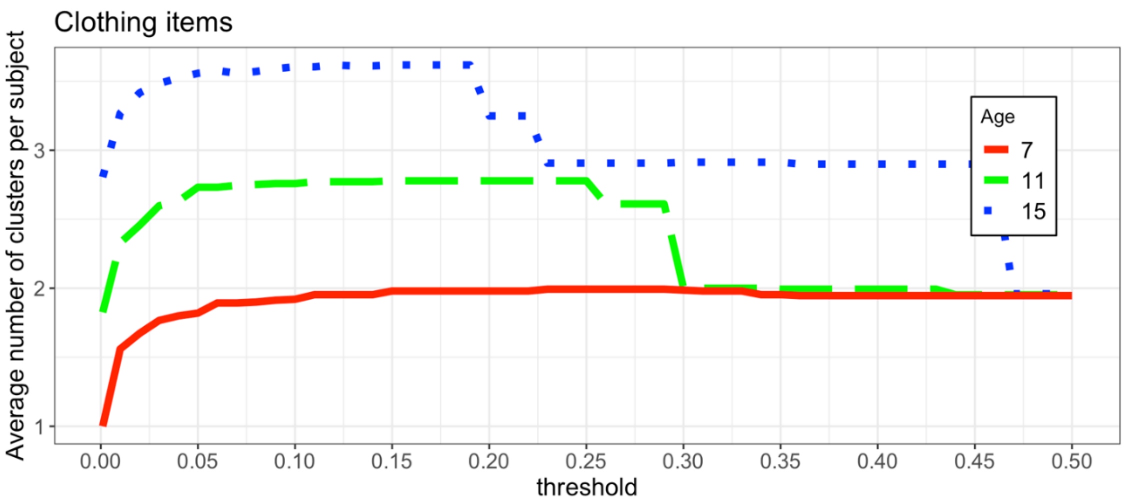


Figure S9. Average number of clusters versus threshold (*t*) for *clothing items*

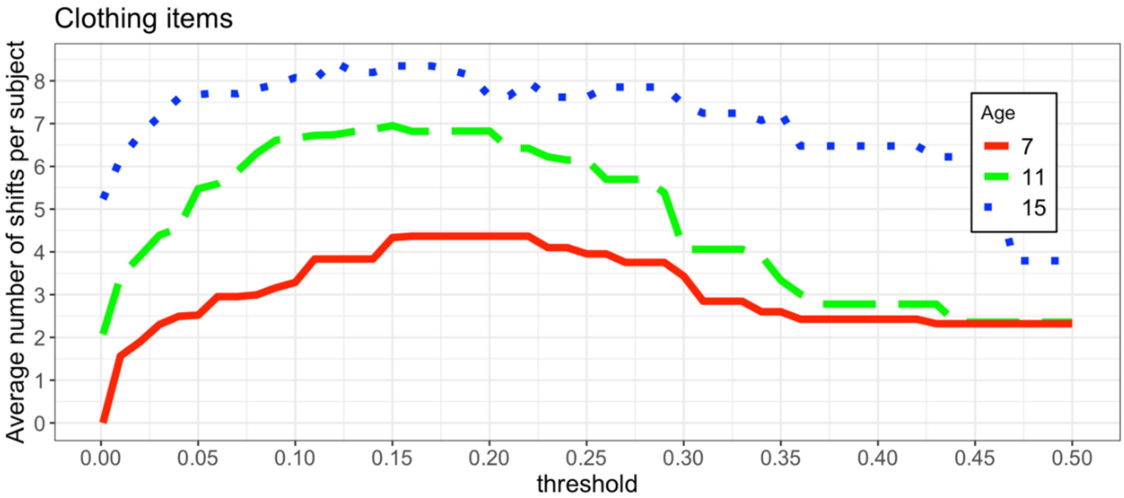


Figure S10. Average number of shifts versus threshold (*t*) for *clothing items*

We can see that the maximum number of clusters and shifts occurs when the threshold is set to approximately .19 considering the three age groups.

Similarly, for *foods*, Figures S11 and S12 show the corresponding plots, where we can see that the threshold has to be set to approximately .02.

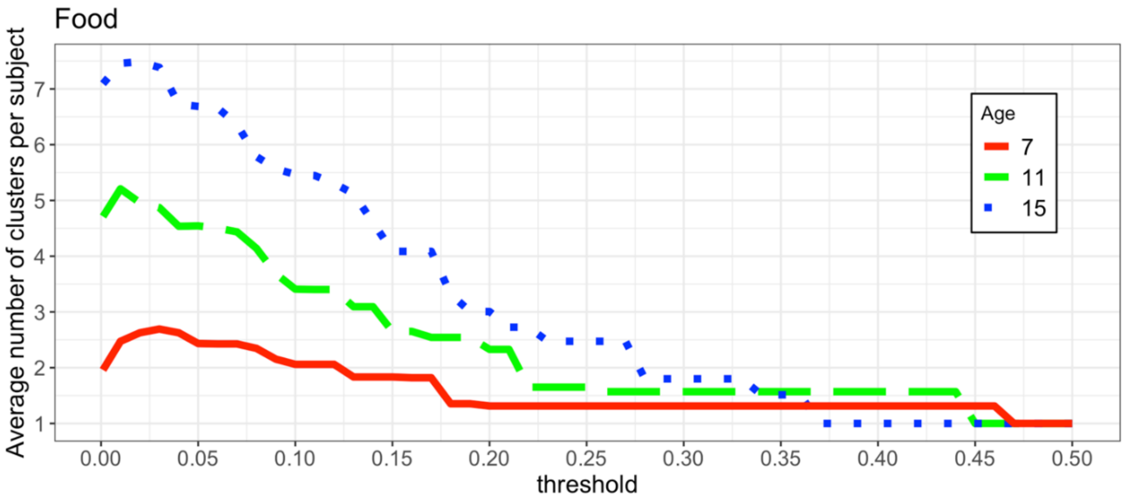


Figure S11. Average number of clusters versus threshold (*t*) for *foods*

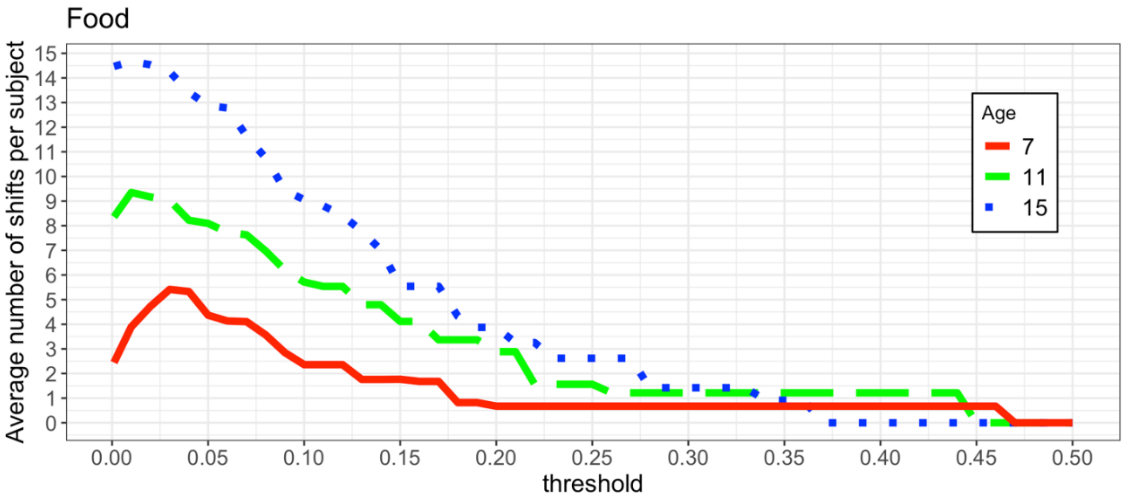
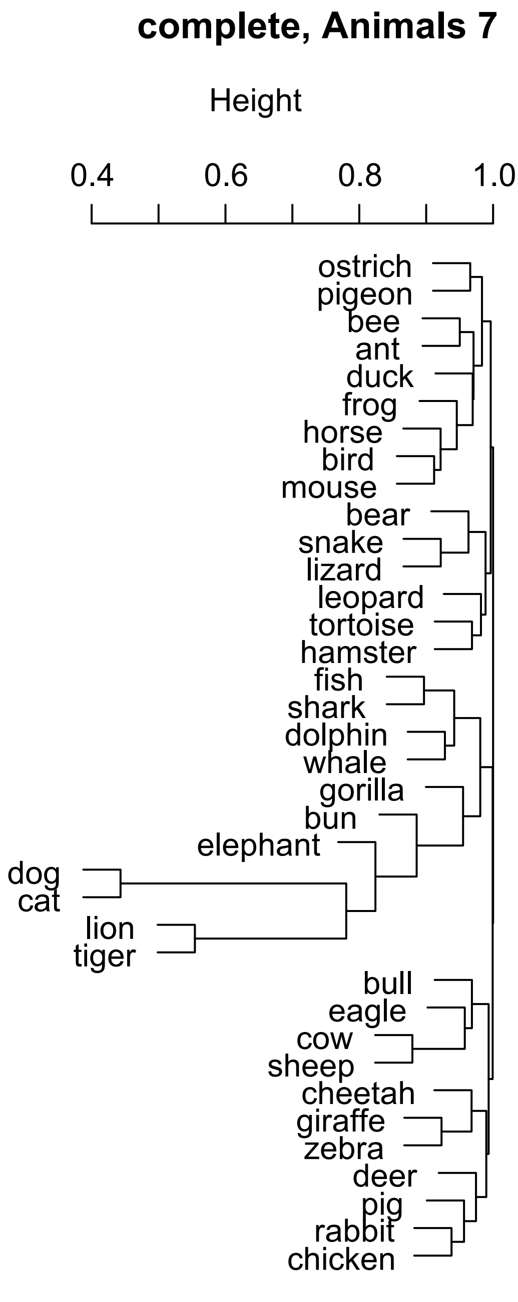
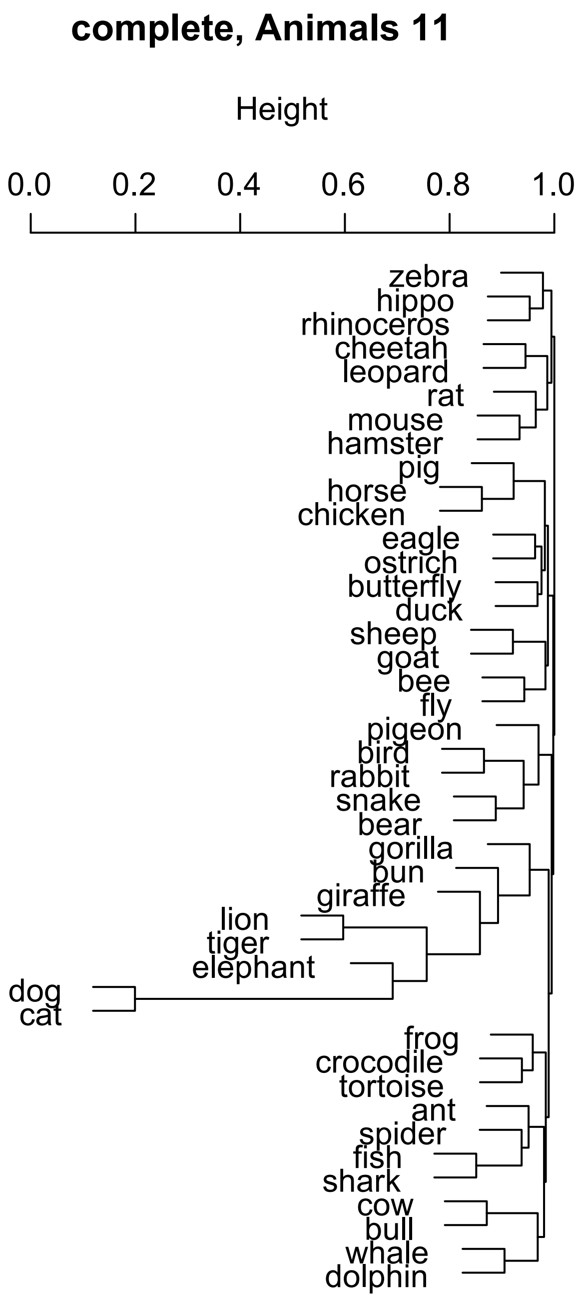


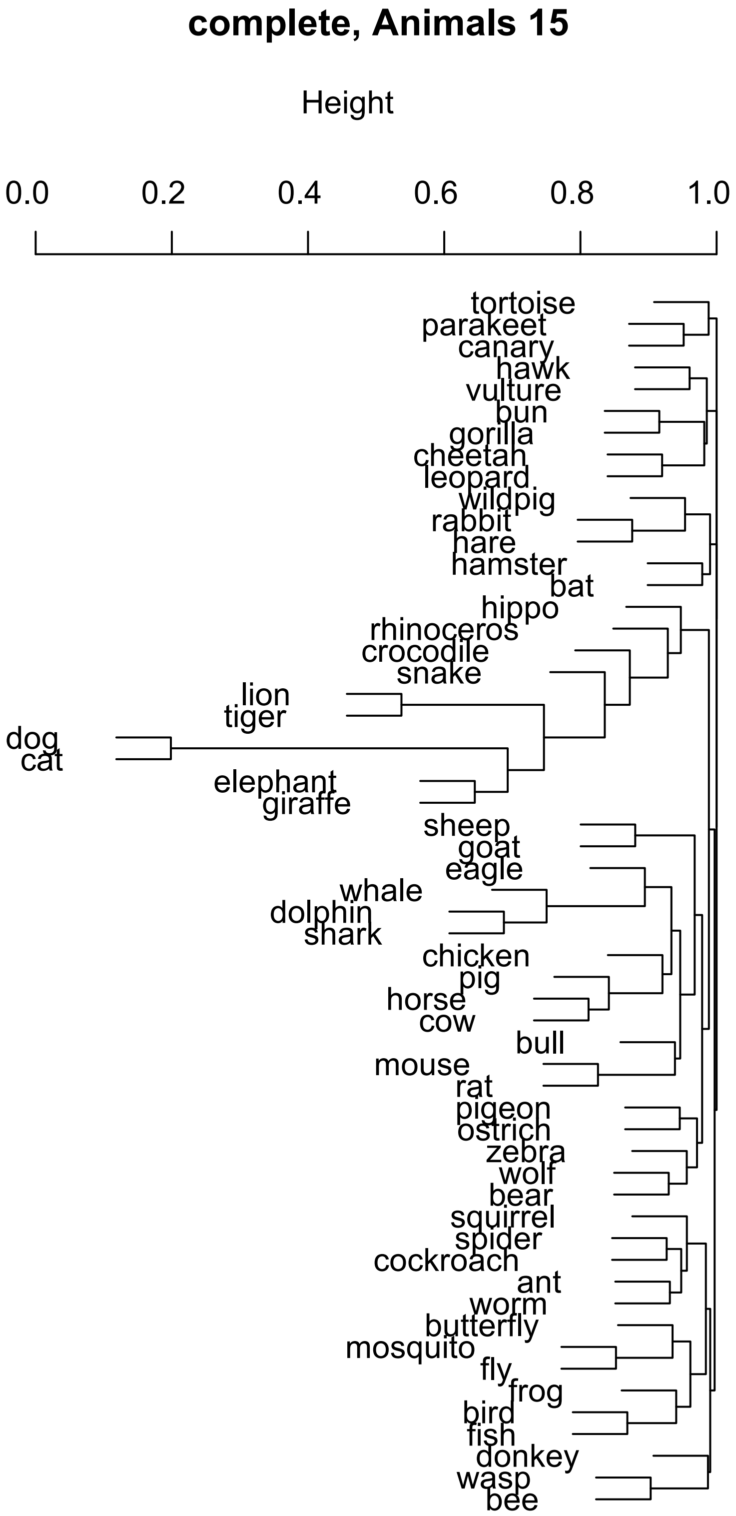
Figure S12. Average number of shifts versus threshold (*t*) for *foods*

Section E. Hierarchical clustering

For the *animals* category, dendrograms show hierarchical clustering for each of the three age groups. On each case, we used the complete linkage method and distance = 1 – similarity. Distances were computed not considering similarities below the threshold used in our analyses (see main text for details). On each case, note that clusters are semantically organized and that the structure increases in complexity with age, with clusters being larger and also more easily interpretable (e.g., the *farm animals* cluster formed by chicken, pig, horse, cow, progressively emerges with age).




complete, Animals 11
Height
0.0
0.2
0.4
0.6
0.8
1.0
zebra
hippo
rhinoceros
cheetah
leopard
rat
mouse
hamster
pig
horse
chicken
eagle
ostrich
butterfly
duck
sheep
goat
bee
fly
pigeon
bird
rabbit
snake
bear
gorilla
bun
giraffe
lion
tiger
elephant
dog
cat
frog
crocodile
tortoise
ant
spider
fish
shark
cow
bull
whale
dolphin


complete, Animals 15
Height
0.0
0.2
0.4
0.6
0.8
1.0
tortoise
parakeet
canary
hawk
vulture
bun
gorilla
cheetah
leopard
wildpig
rabbit
hare
hamster
bat
hippo
rhinoceros
crocodile
snake
lion
tiger
dog
cat
elephant
giraffe
sheep
goat
eagle
whale
dolphin
shark
chicken
pig
horse
cow
bull
mouse
rat
pigeon
ostrich
zebra
wolf
bear
squirrel
spider
cockroach
ant
worm
butterfly
mosquito
fly
frog
bird
fish
donkey
wasp
bee

References

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, *39*(1), 15–22. https://doi.org/10.1080/00221309.1948.9918159

Carriedo, N., Corral, A., Montoro, P. R. P. R., Herrero, L., & Rucián, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology*, *52*(4), 666–678. https://doi.org/10.1037/dev0000091

Carriedo, N., & Rucián, M. (2009). Adaptación para niños de la prueba de amplitud lectora de Daneman y Carpenter (PAL-N). Adaptation of Daneman and Carpenter's Reading Span Test (PAL-N) for children. *Infancia y Aprendizaje / Journal for the Study of Education and Development*, *32*(3), 449–465. https://doi.org/10.1174/021037009788964079

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386–404. https://doi.org/10.1016/0022-0965(82)90054-6

Christ, S. E., Steiner, R. D., Grange, D. K., Abrams, R. A., & White, D. A. (2006). Inhibitory control in children with phenylketonuria. *Developmental Neuropsychology*, *30*(3), 845–864. https://doi.org/10.1207/s15326942dn3003_5

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*(5), 769–786. https://doi.org/10.3758/BF03196772

Daneman, M., & Carpenter, P. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior, 1*, *19*, 450–466. https://doi.org/doi:10.1016/S0022-5371(80)90312-6

De Beni, R., & Palladino, P. (2004). Decline in working memory updating through ageing: Intrusion error analyses. *Memory*, *12*(1), 75–89. https://doi.org/10.1080/09658210244000568

Durston, S., Thomas, K. M., Yang, Y., Uluğ, A. M., Zimmerman, R. D., & Casey, B. J. (2002). A neural basis for the development of inhibitory control. *Developmental Science*, *5*(4). https://doi.org/10.1111/1467-7687.00235

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Fox, C. J., Mueller, S. T., Gray, H. M., Raber, J., & Piper, B. J. (2013). Evaluation of a short-form of the Berg Card Sorting Test. *PloS One*, *8*(5), e63885.

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2009). Individual differences are almost entirely genetic in origin. *Journal of Experimental Psychology*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201.Individual

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Johnstone, S. J., Pleffer, C. B., Barry, R. J., Clarke, A. R., & Smith, J. L. (2005). Development of inhibitory processing during the Go/NoGo task: A behavioral and event-related potential study of children and adults. *Journal of Psychophysiology*, *19*(1), 11–23. https://doi.org/10.1027/0269-8803.19.1.11

Lechuga, M. T., Moreno, V., Pelegrina, S., Gómez-Ariza, C. J., & Bajo, M. T. (2006). Age differences in memory control: Evidence from updating and retrieval-practice tasks. *Acta Psychologica*, *123*(3), 279–298. https://doi.org/10.1016/j.actpsy.2006.01.006

Leon-Carrion, J., García-Orza, J., & Pérez-Santamaría, F. J. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience*, *114*(10), 1291–1311. https://doi.org/10.1080/00207450490476066

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action. *Psychological Review*, *91*, 295–327. https://psycnet.apa.org/record/1994-97487-005

MacLeod, C. M. (2006). The Stroop Task in Cognitive Research. In *Cognitive methods and their application to clinical research.* (pp. 17–40). https://doi.org/10.1037/10870-002

Mayr, S., & Buchner, A. (2007). Negative priming as a memory phenomenon: A review of 20 years of negative priming research. *Journal of Psychology*, *215*(1), 35–51. https://doi.org/10.1027/0044-3409.215.1.35

Mondloch, C. J., Geldart, S., Maurer, D., & de Schonen, S. (2003). Developmental changes in the processing of hierarchical shapes continue into adolescence. *Journal of Experimental Child Psychology*, *84*(1), 20–40. https://doi.org/10.1016/S0022-0965(02)00161-3

Montoro, P. R., Luna, D., & Humphreys, G. W. (2011). Density, connectedness and attentional capture in hierarchical patterns: Evidence from simultanagnosia. *Cortex*, *47*(6), 706–714. https://doi.org/10.1016/j.cortex.2010.05.007

Mueller, S. T. (2011). *PEBL´s Berg Card Sorting Test-64 (PBCST-64). Computer software retrieved from http://pebl.sf.net/battery.html*.

Munro, S., Chau, C., Gazarian, K., & Diamond, A. (2006). Dramatically larger flanker effects (6-fold elevation). *Cognitive Neuroscience Society Annual Meeting*.

Naglieri, J. A., & Das, S. (1997). *Das-Naglieri cognitive assessment system*. Riverside Publishing.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory and Cognition*, *29*(2), 344–354. https://doi.org/10.3758/BF03194929

Pritchard, V. E., & Neumann, E. (2004). Negative priming effects in children engaged in nonspatial tasks: evidence for early development of an intact inhibitory mechanism. *Developmental Psychology*, *40*(2), 191–203. https://doi.org/10.1037/0012-1649.40.2.191

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual section 4: Advanced Progressive Matrices.* Oxford Psychologists Press.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology*, *28*(2), 573–594. https://doi.org/10.1207/s15326942dn2802_2

Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7-10-year-old children. *Cognitive Development*, *22*(4), 474–488. https://doi.org/10.1016/j.cogdev.2007.08.009

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Somsen, R. J. M., Van Der Molen, M. W., Jennings, J. R., & Van Beek, B. (2000). Wisconsin Card Sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica*, *104*(2), 227–257. https://doi.org/10.1016/S0001-6918(00)00030-5

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, *59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Tamnes, C. K., Østby, Y., Walhovd, K. B., Westlye, L. T., Due-Tønnessen, P., & Fjell, A. M. (2010). Neuroanatomical correlates of executive functions in children and adolescents: A magnetic resonance imaging (MRI) study of cortical thickness. *Neuropsychologia*, *48*(9), 2496–2508. https://doi.org/10.1016/j.neuropsychologia.2010.04.024

Tipper, S. P. (1985). The Negative Priming Effect: Inhibitory Priming By Ignored Objects. *The Quartely Journal of Experimental Psychology*, *37A*, 571–590.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. CUP Archive.

Van Den Wildenberg, W. P. M., & Van Der Molen, M. W. (2004). Additive factors analysis of inhibitory processing in the stop-signal paradigm. *Brain and Cognition*, *56*(2 SPEC. ISS.), 253–266. https://doi.org/10.1016/j.bandc.2004.06.006

Verbruggen, F., Aron, A. R., Band, G. P. H., Beste, C., Bissett, P. G., Brockett, A. T., Brown, J. W., Chamberlain, S. R., Chambers, C. D., Colonius, H., Colzato, L. S., Corneil, B. D., Coxon, J. P., Dupuis, A., Eagle, D. M., Garavan, H., Greenhouse, I., Heathcote, A., Huster, R. J., … Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *ELife*, *8*. https://doi.org/10.7554/ELIFE.46323/ELIFE_46323_PDF.PDF

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods*, *40*(2), 479–483. https://doi.org/10.3758/BRM.40.2.479

Waszak, F., Li, S. C., & Hommel, B. (2010). The development of attentional networks: Cross-Sectional findings from a life span sample. *Developmental Psychology*, *46*(2), 337–349. https://doi.org/10.1037/a0018541

Wilcox, R. R., & Keselman, H. J. (2003). Modem robust data analysis methods: Measures of central tendency. *Psychological Methods*, *8*(3), 254–274. https://doi.org/10.1037/1082-989X.8.3.254