LUNDS UNIVERSITET

**Department of Psychology**

# A Snapshot into Developmental Differences in Attention and Cognitive Control: An Eye Tracking Study

**Molly Sterlacci**

Master's Thesis (30 hp)
Spring/Summer 2017

Supervisors: Elia Psouni
Roger Johansson

Abstract

The present study reveals how the presence of relevant and irrelevant spoken language affects the processing for visual information, and if this is mediated by executive function processes that develop from childhood ($N$ = 10, age 4-10) to adolescence ($N$ = 16, age 14-17) and through adulthood ($N$ = 16, age 21-28). Participants engaged in a task in which they looked at visual stimuli organized within Areas of Interest (AOI's) with no audio (1), with audio to facilitate memory (2), and distracting audio unrelated to visual stimuli (3) and were then asked to remember the visual stimuli they looked at, all while connected to an eye tracker. Results showed that there was no main effect of distraction or facilitation conditions on overall memory performance. However, there was a significant difference in memory performance between age-groups, with the children unexpectedly performing the best, followed by adults and then adolescents. Also, memory performance outcomes were not mediated by executive function. Dwell times, fixation counts, and fixation times between matching audio and visual information and non-matching audio and visual information were not significantly different. However, dwell time and fixation time were significantly affected by age, with children having lower times on both non-matches and matches. Limitations of the study and areas for future research are discussed.

*Keywords:* attention, distraction, executive function, cognitive control, eye tracking, memory

# A Snapshot into Developmental Differences in Attention and Cognitive Control: An Eye Tracking Study

Selective attention is one mechanism that selects goal-relevant stimuli for further processing while inhibiting the processing of goal-irrelevant stimuli. While selective attention is adaptive in order to prevent sensory overload, it is not always successful (Gatti & Egeth, 1978). Lavie and colleagues (Lavie, 1995) have suggested that one factor in determining the success of selective attention is the perceptual load (PL) of the task, whereby a low PL does not use all resources for target processing, so leftover resources can be used for distractor processing. Conversely, when the PL is high, all resources are utilized on target processing, and distractors are not processed. Also, a high load on ‘frontal’ cognitive control such as working memory processes increases distractor processing. However, much of the work supporting perceptual load theory (PLT) comes from methods using the flanker, Stroop, Simon, and prime-probe tasks (Eriksen & Eriksen, 1974). Furthermore, most of this work has been focused on the use of visual stimuli. Yet, relatively little is known about to what extent top-down executive function capabilities prevent distraction by salient but irrelevant stimuli within a cross-modal context.

The current study uses eye tracking to investigate the processing of task-relevant facilitators and task-irrelevant distractors in a framework that integrates attention research with executive function research. Specifically, subjects are tested on their retention of previously encoded visual information held in working memory when the setting presents a low perceptual load but a high cognitive load. The aim of the current study is to examine whether the presence of relevant and irrelevant spoken language affects the processing for visual information, and if this is mediated by executive function processes that develop from childhood through adolescence into adulthood. Furthermore, it aims to investigate how the gaze patterns interact with the spoken auditory input. Specifically, when the audio information and the visual information in the AOI are a match, is there a shorter dwell time, fixation time and a lower fixation count on that AOI compared to when the audio information and the AOI are not a match? To accomplish this, I will investigate the proportion of dwell time, fixation time and fixation counts on matching AOI's compared with those on non-matching AOI's. The study is based on a mixed, between-group combined with within-subject design, comparing adolescents' responses to those of a child group and an adult group, across conditions involving different combinations of attentional distractors.

It is important to understand these behaviors as research shows that low levels of executive function are associated with clinical-level ADHD symptoms (Campbell & von Stauffenberg, 2009). Furthermore, ADHD identified in early childhood predicts an increased likelihood of functional impairment into adolescence (Lahey et al., 2016). Research estimates that between 50% and 80% of children diagnosed with ADHD continue to experience symptoms and functional impairment through adulthood, with the most common symptoms including distractability, restlessness, and impulsiveness (Barkley, 1998).

## Attention

Attention is invaluable in everyday life. Being able to attend to specific stimuli is necessary in order to avoid getting hit by cars as we cross the street, to search for missing objects, and to perform two tasks at the same time. During every waking moment, humans are flooded with stimuli from multiple sensory modalities, including from non-sensory sources such as memories, emotions or trains of thought occurring to us as we strive to solve problems we encounter or make necessary decisions. In order to successfully navigate the real world, our neural resources must make it possible to attend to the most salient aspects of information, the ones likely to matter from moment to moment. The ability to remain focused on a task is vital for coherent cognitive functioning, especially when there might be potential interference from distractors that are irrelevant for the task.

While attention can be defined in a multitude of ways, it typically refers to selectivity of processing, as was emphasized in the 19$^{th}$ century by William James when he said that "attention is…the taking into possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence" (James, 1890, p. 403-404). James distinguished between active and passive modes of attention, where active attention occurs when controlled in a top-down way by the individual's goals or expectations and passive attention occurs when controlled in a bottom-up way by external stimuli. This distinction, which has also been referred to as controlled versus automatic processing, is still a valid and important construct in the current research on visual attention.

A major goal within the field of attention has been to understand the ways in which attention affects the processing of sensory input. Broadbent (1958) proposed a selective filter theory, asserting that attentional selection can occur very early in processing. More specifically, he proposed that sensory input can be funnelled at early sensory processing stages based on the fundamental physical properties of the stimulus, thus determining which information will proceed to higher levels of analysis. However, this early processing theory ran counter to evidence that some information, such as one's name, can break through the unattended auditory input channel and reach the level of semantic analysis.

Such findings led researchers to propose a late-selection model, in which attentional filtering occurs relatively late in stimulus processing. According to such theories, all sensory stimuli are fully processed, in terms of both their physical properties and their possible higher-level meaning (Deutch & Deutch, 1963; Norman, 1968). In this view, it is not until after the higher-level processing is complete that attention exerts its influence and determines which input information entered consciousness or influences behavior.

Treisman (1960) proposed an alternative theory, an adaptable early-filtering system that could attenuate the inputs from concurrent sensory channels in a flexible way. According to this theory, stimulus analysis proceeds systematically through a hierarchy, starting with analysis based on physical cues, syllable pattern and specific words, and moving on to analysis based on grammatical structure and meaning. If there is insufficient processing capacity to permit full stimulus analysis, tests towards the top of the hierarchy are omitted. In this model, some unattended semantic information, such as one's name, can reach higher levels of conscious awareness.

Subsequently, researchers concluded that an even better model to explain the various findings was one in which incoming information could be filtered at different levels of processing, depending on the needs of the task (Lavie, 2005). So, in some situations, sensory information might be filtered out or selected at relatively early processing stages according to basic physical characteristics such as color or pitch. Under other circumstances, more complex aspects of the information in a stimulus, such as its semantic content, would be required to make the selection, which would then occur at a later stage.

While these theories have contributed to our understanding of voluntary attentional tasks, where subjects follow instructions to direct their visual attention towards objects in the environment, people are often distracted by task-irrelevant stimuli. Daily life provides numerous examples: a bug flying around might distract a reader, an attractive billboard can distract a driver, and so forth. What we see, hear, feel and remember depends not only on the information entering our senses, but also on the aspects of this information we choose to attend to. In the laboratory, controversy has arisen over the extent to which distractor processing can be prevented. Whether focusing attention on task-relevant stimuli can exclude distractors from early perceptual processing (an 'early' selection effect) or can only prevent distractors from controlling behavior and memory (a 'late' selection effect) has fuelled a longstanding debate between early and late-selection views of attention.

However, simply instructing subjects to ignore goal-irrelevant stimuli is not sufficient to prevent that such goal-irrelevant stimuli are processed. Research indicates that distractor processing depends critically on the level and type of load involved in the processing. Specifically, perceptual load theory (PLT) maintains that perceptual resources have a limited capacity, with a requisite use of all available resources. Thus, when the perceptual load (PL) of the task is low, not all resources are utilized on target processing, and thus spill over to distractor processing. On the other hand, when the PL is high, all resources are used on target processing meaning that distractors are not processed, and thus interference with target processing is reduced (Lavie, 1995). Furthermore, a high load on 'frontal' cognitive control such as working memory processes increases distractor processing (Lavie, 2005).

While PLT is well established in the visual domain, less is known about processing between modalities. Early seminal work by Allport and colleagues (1972) demonstrated that a complex auditory task (i.e. auditory shadowing) can be performed at the same time as a demanding visual task (i.e. sight reading music), suggesting a limited effect of processing load across modalities. However, contrasting evidence was provided by Tellinghuisen and Nowak (2003), who used a version of the response-competition task adapted to a cross-modal context. When peripheral letter distractors were presented auditorily during search for visual letter targets, they, unlike visual distractors, filter into further processing stages, causing interference. So, it is unclear how auditory distraction can affect participants during a memory task, and if there is an effect across modalities.

Several studies have shown that attention plays a key role in development. The evaluation of attention in children is important because of its implications on learning, social functioning, and academic achievement (Spira & Fischel, 2005). Abnormal attention development is a symptom of attention deficit hyperactivity disorder (ADHD) concerning functional and structural brain pathological changes (Wilcutt et al., 2005). Attention-deficit/hyperactivity disorder (ADHD) is a childhood-onset condition marked by patterns of inattention and/or impulsivity-hyperactivity that often persist into later life. ADHD symptoms have been shown to be related to future academic and employment difficulties in early adulthood (Pardini, D.A., & Fite, P.J., 2010; Mordre et al., 2012). Moreover, adults with ADHD have been shown to perform worse than controls on verbal long-term memory and memory acquisition (Skodzik et al., 2013). The role of attention in promoting positive outcomes for learning and academic achievement is hence important to understand.

While children and adults perform similarly on visually orienting to single item objects in a barren visual field, it is unclear when individual's goals can begin to control what is perceived. It has been shown that sudden peripheral changes and single items in an otherwise empty visual field can capture attention and trigger orienting responses toward their location comparably in six and eight year olds as in adults (Enns & Brodeur, 1989). However, research shows that adults exert top-down control to avoid distraction by salient yet irrelevant stimuli, but this capacity is much less developed in children. For instance, Gaspelin and colleagues showed that children at 4.2 years of age were more vulnerable to be captured by irrelevant stimuli than adults of 21.5 years (Gaspelin, Margett-Jordan, & Ruthruff, 2015). Nevertheless, it is unclear when this top-down control competence develops between childhood and adulthood. No investigations of adolescents' attentional capacity on a task which has low perceptual load combined with a high cognitive load has been presented to the author's knowledge, leaving the question of whether adolescents can manage distraction by irrelevant stimuli like adults unanswered.

Finally, it has been shown that attention can impact memory performance. For instance, Schmitter-Edgecombe (1996) demonstrated that undergraduate students who performed a word-rating task in a full attention condition performed significantly better on an explicit memory task using cued recall than subjects in the divided attention condition. However, to the author's knowledge, no research has investigated how explicit memory performance is impacted by an attentionally demanding task across modalities.

**Executive Function**

Definitions of executive function (EF) differ widely. One approach considers EF as a higher order cognitive mechanism or ability. Baddeley's influential model comprises three interacting component parts; an executive controller and two subservient systems (Baddeley & Hitch, 1994). The executive controller functions as an attentional control system. The two subservient systems, one underpinning verbal processing and the other underpinning visual processing are themselves subdivided. In the verbal system, a passive phonological store is maintained by an active phonological loop, which is able to rehearse the material in the passive store. The visual working memory system has traditionally been thought of as having a similar architecture with a passive visual store being maintained by an active store, which codes in terms of movement over space.

Baddeley attempts further to specify the central executive component of working memory, and analyze its component functions. He divides executive function into four functions: the capacity to manage performance on two separate tasks, the capacity to switch retrieval strategies, the capacity to selectively attend to one stimulus and inhibit the disrupting effect of another, and the capacity to hold and control information in long-term memory, as reflected in measures of working memory span (Baddeley, 1996). More recent work has identified the functions of EF to include: maintaining information in working memory, inhibition of prepotent responses, planning, rule use, error correction, recognition of and resistance to interference, mental flexibility, and shifting and sustaining of attention in carrying out a goal (Carlson, Faja, & Beck, 2015). For the purpose of this study, I will define EF as the ability to effectively maintain task goals in the face of distraction, thereby exerting cognitive control over the contents of thought and action.

Executive function is primarily associated with the prefrontal cortex and most knowledge about it comes from neuroimaging data and clinical observations of children and adults with brain lesions or medical conditions known to affect prefrontal functioning. Several studies using various measures of self-control in children have suggested that EF plays an important role in cognitive and social development, such as theory of mind, emotion regulation, moral conduct, and school readiness (Blair & Razza, 2007; Carlson & Wang, 2007; Carlson, Mandell, & Williams, 2004; Shoda, Mischel, & Peake, 1990). For instance, Kochanska and colleagues found that early differences in attention and effortful control predicted later social functioning, including moral conduct (Kochanska, et al., 1997). Conversely, disruptions in EF are implicated in a number of childhood disorders and are associated with poor social and academic adjustment (Casey, Tottenham, & Fossella, 2002; Hughes & Ensor, 2011), persisting even into adulthood (Casey et al., 2011; Moffitt et al., 2011). Therefore, the ability to maintain cognitive control in the face of distraction is an important life skill.

Although some research has suggested that executive control develops throughout early childhood (Cepeda, Kramer, & Gonzalez de Sather, 2001; B. R. Williams, Ponesse, Schachar, Logan, & Tannock, 1999), it is unclear whether this empirical generalization applies specifically to the control of visual attention while attending to an auditory stimulus, which has high practical importance.

## Visual Attention, Eye Movements and Language

To obtain a complete understanding of the role of attention in visual cognition, it is necessary to understand eye movements. The two main eye movements are fixations and saccades. During a fixation, the eye is relatively stable and pauses on an area of interest to process visual information. Fixations last anywhere from tens of milliseconds up to several seconds (Rayner, 2009). It is generally considered that a fixation measures attention to that position. The rapid motion of the eye from one fixation to another is called a saccade. Saccades are very fast, typically taking beteween 30-80 milliseconds (ms) to complete (Holmqvist et al., 2011).

There is ample evidence that fixation placement during scene viewing is strongly affected by cognitive factors. According to the visual saliency hypothesis, fixation points are selected based on image properties generated in a bottom-up fashion from the current scene. In this way, gaze control is a reaction to the visual properties of the stimulus in front of the viewer. On the other hand, according to the cognitive control hypothesis, fixation points are selected based on the needs of the cognitive system in relation to the current task (Henderson et al., 2007). In this way, eye movements are primarily controlled by task goals interacting with a memory for similar viewing experiences and a semantic interpretation of the environment.

The assumption that underlies most eye tracking research is that what we look at closely corresponds to what we are thinking about (Yarbus, 1967). By tracking a participant’s eye movements, it is possible to measure precisely what information is available for processing. Eye tracking has been used to provide insights into attention abilities in populations including children and adults alike.

Language, too, influences what we look at in a visual scene. Research by Cooper (1974) and by Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995) demonstrated how visual attention towards the external world can be mediated by the unfolding spoken language. As sentences unfold within visual contexts depicting the objects mentioned in those sentences, the eyes look towards those objects. This phenomenon, known as the visual world paradigm, provides information about the way language users integrate linguistic information with information derived from the visual environment. Further research indicated that even as words unfold, the eyes begin to move towards the corresponding objects (Altmann & Kamide, 2007). For instance, it was demonstrated that when a sentence such as '*the boy will eat the cake*' is heard in the context of a scene depicting a boy and a cake (and other things), the eyes will begin to look towards the cake during the verb '*eat*'(Altmann & Kamide, 1999). Therefore, it is important to further investigate the interplay between language and visual attention.

**The Current Study**

The current study uses eye tracking to investigate the processing of task-relevant facilitators and task-irrelevant distractors in a framework that integrates attention research with executive function research. Subjects are tested on their retention of previously encoded visual information held in working memory when the setting presents a low perceptual load but a high cognitive load. I want to know how the presence of relevant and irrelevant spoken language affects the processing and memory for visual information, and if this is mediated by executive function processes. I want to understand how the gaze patterns interact with the spoken auditory input. Specifically, when the audio information and the visual information in the AOI are a match, is there a shorter dwell time, fixation time and a lower fixation count compared to when the audio information and the AOI are not a match? To accomplish this, I will investigate the proportion of dwell time on a matching AOI compared with dwell time on a non-matching AOI. Furthermore, the study will test the relationship between language and eye movements, as has been demonstrated in the visual world paradigm.

**Hypotheses**

Hypothesis 1: When participants are exposed to three separate conditions (control with no verbal utterances, matching audio and visual information, and non-matching audio and visual information), there will be an effect on the memory performance, such that the highest performing to lowest memory performance order is: matching, control, non-matching.

Hypothesis 2: Age group will significantly impact memory performance, such that there will be a significant difference between the children and adolescents and adults both, but not between the adolescents and adults. The adolescents and adults will perform similarly and better than the children.

Hypothesis 3: Executive function is critical for performing well under conditions of distraction. Hence, memory performance outcomes will be mediated by scores of EF when the audio information is not matching any of the images for later memory recall.

Hypothesis 4: When there is a non-match between audio and visual information during encoding, it will lead to longer dwell times, a higher fixation count, and longer fixation times on non-matched AOI's than when there is a match between audio and visual information to AOI.

Hypothesis 5: Age group will significantly impact dwell times, fixation times and fixation counts on both audio/visual matched and audio/visual non-matched AOI's, with shorter times and fewer counts for the children group, and longer times and more counts for the adolescents and adults.

# Method

The study was based on a mixed, between-group combined with within-subject design, comparing adolescents' responses to those of a child group and an adult group, across conditions involving different combinations of attentional distractors.

## Participants and Sampling

A total of 42 individuals (23 females), consisting of 10 children whose mean age was 6.4 years (*SD* = 1.6), 16 adolescents whose mean age was 15.2 years (*SD* = 1.0), and 16 adults whose mean age was 23.9 (*SD* = 2.1) took part in the study. However, one of the adolescent participants' eye-tracking data was lost due to technical problems. Thus, 41 participants were included in the analysis. The children were between 4 years and 10 months to 10 years and 1 month old. The adolescents were between 14 years and 9 months to 17 years old. Finally, the adults were between 21 and 28 years old. All participants were native Swedish speakers. All participants reported normal or corrected-to-normal visual acuity. Two participants (5%) wore glasses.

## Psychometric assessment

I administered a psychometric assessment for each participant. Specifically, the children performed two subtests from the Swedish version of the Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV, Wechsler, 2012) to assess cognitive development. The first test assessed general vocabulary knowledge (Word Recognition) and the second assessed verbal intelligence (Information). The Word Recognition test gives scores from 0 to 31 while the Information test gives scores between 0 and 29. The WPPSI-IV is intended for children aged 2 years 6 months to 7 years 7 months. The Wechsler Preschool and Primary Scale of Intelligence is a well-established scale and has received high consistency. The internal consistency coefficients range from 0.85 to 0.96, overall inter-scorer agreement ranges from .98 to .99, and high uncorrected stability coefficients of .78 to .88 have been reported on a short-term test-retest of stability (Wechsler, 2012).

To assess executive function in the children's group, I administered two tests: Dimensional Change Card Sort (DCCS; Zelazo, 2006) and the Day-Night Stroop-like task (Gerstadt, Hong and Diamond, 1994). The DCCS gives scores from 0 to 12 while the Day-Night task gives scores between 0 and 16. Performance on the DCCS provides an index of the development of executive function, and it is impaired in children with developmental disorders, such as ADHD and autism (Zelazo, 2006). The Day-Night Stroop-like task necessitates inhibitory control of action as well as learning and remembering two rules (Gerstadt, Hong and Diamond, 1994).

I administered similar age-adapted tasks for the adolescent and adult population, namely two subtests from the Swedish version of the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, 2008). The first test assessed general vocabulary knowledge (Vocabulary) and the second assessed verbal intelligence (Information). The Vocabulary test gives scores between 0 and 57 while the Information test gives scores between 0 and 26. The WAIS-IV is intended for use among adolescents to adults, aged 16-90 years. The Wechsler Adult Intelligence Scale is a well-established scale and has fairly high consistency. Over a two-to-twelve-week period, the test-retest reliabilities ranged from 0.70 to 0.90 (Wechsler, 2008). High inter-scorer coefficients (above .90) have also been reported. The WAIS correlated highly with the Stanford-Binet IV test (0.88) and had high concordance with various measures: memory, language, attention, and cognitive ability (Wechsler, 2008).

In addition, to assess executive control abilities, I administered a Stroop task to adolescents and adults. As it was not possible to find a publically available Swedish version for the Stroop task, I programmed a Swedish version of the task myself using Psychopy and

Python script. The accuracy scores ranged from 0 to 30, and a mean response time was also recorded. In total, the psychometric assessment part took approximately 15 minutes per participant.

**Task Design and Materials**

Participants engaged in a task in which they looked at visual stimuli while listening to audio stimuli and were then asked to remember the visual stimuli they looked at. First, participants were presented with four plain white doors that corresponded to four locations on the screen. Then, the white doors opened to initiate the encoding stage of the experiment, whereby the same four locations now corresponded to colored backgrounds with objects in each one. Visual stimulus targets consisted of images of commonplace objects (e.g. balloon, car, elephant, pear) with colored backgrounds and arranged in the form of a 2 X 2 matrix, with the central position of each cell occupied by a different stimulus target. An example is seen in Figure 1. The uniform arrangement of target elements into 2 X 2 arrays of equal dimensions was used to simplify scoring and calibration procedures.

As can be seen in the example in Figure 1, the four images on the screen also corresponded to various background colors, namely blue, red, green, and yellow. The reason for this is so that each image had its own unique color and location. The combinations of colors and locations for the images changed from trial to trial, so that participants could not associate, for example, an image of a strawberry as always having a red background, or always being in the top right corner.

It is generally agreed that visuospatial processing is accomplished in distinct neuroanatomic pathways. Namely, the dorsal route (the “where” pathway) which conveys location and motion information and the ventral route (the “what” pathway) which conveys color and form information. The where pathway is thought to be responsible for processing spatial relationships while the what pathway is responsible for object identification (Ungerleider & Mishkin, 1982). A central question in developmental psychology is how a child acquires knowledge about the surrounding world. Behavioral research investigating this has distinguished the role of dorsal and ventral visual streams in learning to “know what” and “know how” about objects, but these studies have not definitively resolved how these functions develop (Braddick, Atkinson, & Innocenti, 2011). Hence, ongoing research by Elsinga and Psouni is investigating cueing towards the two visual pathways when children are in a memory task to better understand the way these systems interact through development. So, this was the motivation for incorporating cues both toward location (the “where” pathway cue) and color (the “what” pathway cue). However, the discussion of the implications of such cues on memory performance and gaze patterns is not within the scope of this paper and can be found elswewhere.

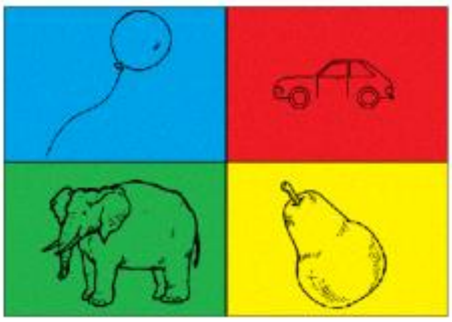

*Figure 1.* Examples of stimuli (adapted from ongoing research by Elsinga and Psouni).

Participants were instructed to look at the pictures and try to remember them. However, they were given no explicit instruction regarding the auditory stimuli. The images were present for 9000 milliseconds, during which time auditory stimuli were also presented. This typically left 2 s after the end of each sentence before the encoding phase was terminated. Auditory stimulus materials consisted of tape-recordings of full sentences in a normal speaking rate by a male native speaker of standard Swedish. Each participant was exposed to three separate conditions throughout the experiment: no-verbal utterance (control), verbal utterance – matching, and verbal utterance – non-matching. The first condition was intended as a baseline, in which participants had no auditory input to process. In the verbal utterance – matching (second condition), one of the objects in the visual display was the referent in the verbal utterance. As an example, if the object was a bucket, the sentence was ‘The bucket is full of water.’ In the verbal utterance – non-matching (third condition), none of the objects in the visual display were referred to. As an example, when the four objects presented were king, hamburger, shoe, and umbrella, the sentence was ‘The stereo is playing loud music.’ The second condition and third conditions were intended as experimental manipulations that would facilitate (second) and distract from (third) the task at hand. Participants did not have to recall the spoken phrase in the second and third conditions. Thus, the visual information was task-relevant and the auditory information task-irrelevant. A short distraction task followed, in which the goal was to identify and verbally state the location of the cartoon figure that moved from one object to another object. An example is seen in Figure 2.

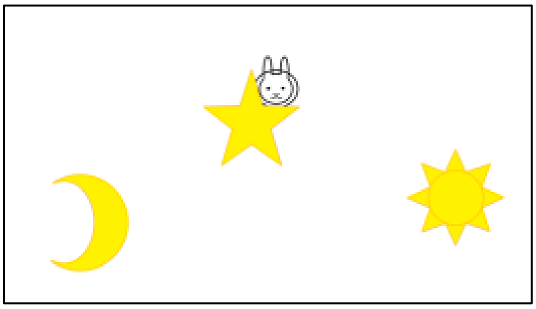

*Figure 2.* Example of distraction task (adapted from ongoing research by Elsinga and Psouni).

After the distraction task was the recall phase. During this phase, participants were presented with a location or color cue, and asked to recall the image which was in that location or the image which was matched with that color. In the case of a location cue, a white door was presented in the quadrant matching the recall image during encoding. In the case of a color cue, a colored door was presented in the central position of the screen, matching the background color for the recall image during encoding. Participants were asked “Do you remember what was behind the door?” before being allowed to verbally answer the name of the image corresponding to that color/location if they could remember it. I coded their verbal responses as either incorrect or correct. There were two separate cues presented during each trial (two responses coded on each trial), and these cues were either both to location or both to color.

During pilot testing with children, I realized that the task was overly challenging for the children’s group, and they had difficulty maintaining focus and motivation during the trials. Therefore, I decided to offer two cues for the children’s group during the recall phase, such that after asking whether they remembered what was behind the door, they were given one cue which was the correct answer and another which was an image from the same trial (e.g. “was it a king or a shoe?”). The order of the two cues was semi-randomized over trials, so that the correct answer was not always in the first position or second position. However, the adolescent and adult groups were not offered these cues. To ensure that the mix of colors, locations, images, cues and order of conditions was semi-randomized over all trials, I utilized a random number generator tool (www.random.org). The result of this semi-randomization was that the images were repeated no more than twice during the experiment, nor were they repeated as a target image during recall more than once. There were 12 trials in the experiment, determined after a pilot test performed in association with research by Elsinga and Psouni. Of the 12 trials, four were the control condition, four were the matching condition, and four were the non-matching condition. There were 12 different versions of the experiment with varying combinations regarding the order of conditions, images, colors, locations, and cues. The entire experiment lasted approximately 35 minutes per participant.

**Equipment and Data Acquisition**

*Equipment.* Participants were seated 600 mm in front of a 480×300 mm monitor with a resolution of 1680×1050 pixels. The experiment was implemented using Psychopy and Python scripts. Stimulus material was presented using Psychopy, while participants' eye movements were recorded using an SMI iView RED250 eye tracker (Teltow, Germany) recording binocularly at 250 Hz. Data were recorded with the iView X 2.5 software (SensoMotoric Instruments, Teltow, Germany). I also chose to use a chin-rest to cut down on potential head movement by participants, which would threaten the quality of the data. During pilot testing, I discovered that the chin-rest was intimidating to the children, so I decorated it a bit to create a space story, whereby they sat in a rocket shaped tent ("space ship") and looked out of a "space window," which was really a piece of decorated cardboard with a large hole cut in the spot where the head sat on the chin rest. This way, it was more fun for the children, and they sat their heads on the chin-rest without being intimdated by the look of the chin-rest contraption. In addition, the whole experiment was video recorded in case there were unclear responses that needed to be re-evaluated. BeGaze 3.4 and SPSS 22 were used for analysis.

*Calibration:* I performed a two-point calibration with validation. If the deviation of the validation values was above 2.0 visual degrees, I repeated the process until a value below 2.0 was reached. I chose a two-point calibration because the Areas of Interest (AOI's) in the experiment were large enough that minute precision of the eye tracking points was not needed. Additionally, the two-point calibration was easier to perform with the children's age group.

*Pictures.* The 24 images which were shown in the experiment were line drawings of common everyday objects, which would be easily recognizable to children and adults alike. The images were drawn from various sources, including ten from a picture series by Snodgrass and Vanderwart (1980), eight from the CRL International Picture-Naming Project (IPNP), and six others from an internet search and had a comparable style. The pictures from the CRL International Picture-Naming Project are publically available for downloading on their website (www.crl.csd.edu). They were all chosen to be visually similar in terms of size and shape as well as in word frequency.

**Procedure**

*Recruitment:* I personally recruited participants through marketing the study to local elementary schools, high-schools, as well as through local places of business, and participants were asked to come to the Institution of Psychology at the local University to participate in the study. The experiment took place in a university laboratory. Participants were told that the experiment was interested in investigating susceptibility to distraction. All materials used were in Swedish. Participants included Swedish native-speaking children (aged 4-10), adolescents (aged 14-17), and adults (aged 21-28).

*Testing:* Before the experiment started, participants reported their age, gender, and whether they have normal or corrected-to-normal visual acuity. Participants performed the psychometric tests of knowledge, information, and executive function. Participants then performed the memory task in which they were asked to look at the images on the computer screen and try to remember them. For the children's group, the experimental computer task was broken up into two segments, with the DCCS test being performed in between the two halves, to not fatigue the young participants for too long with one task.

*Debriefing:* Once the complete experiment was finished, participants (and if applicable, their caregivers) were debriefed on the nature of the study and offered a chance to ask any questions. Finally, participants were rewarded for their participation. Adolescent and adult participants were rewarded with 1 movie ticket while the children were rewarded with a small toy of their choice.

## Ethical Considerations

An application to the regional Ethical Review Board in Lund (EPN) was submitted for ethical clearance of the project in agreement with the Act regarding the Ethical Review of Research Involving Humans. The project was regarded as not involving risk of harm to participants (either physical or mental), nor was there a concern for the handling of personal information. Therefore, the project did not necessitate formal ethical clearance.

All participants gave their informed consent and in the case of the participant being younger than 15 years of age, informed consent was obtained through parental guardian. The researcher informed the participants that they had the right to withdraw from the study at any given point without any consequences. The researcher respected confidentiality and anonymity. No one could link a given participant's performance on any one task to the other tasks, as the participants were identified with a key code, so the researcher could identify the corresponding data, but did not know whose responses they were. Only the researcher and supervisors had access to the collected data. Sensitive information was not asked of the participants. At the end of the experiment, participants were debriefed. There was minimal risk of harm to participants. No material used in the study was directly upsetting to the participants, nor was deception involved.

## Analysis Plan and Data Preparation

*Missing Data and data cleaning:* One adolescents' eye tracking data was lost due to technical issues. Five trials had data loss over 50%, and were removed from the analysis. All in all, there was little missing data. All participants completed all trials of the experiment.

*Eye Tracking Variables:* The raw eye tracking data was transformed into fixations and saccades based on AOI's by eye tracking analysis software BeGaze version 3.7. This file was exported from BeGaze version 3.7 as an Excel document, and included a line for each sample recorded by the eye tracker. To gather the critical parts of the recordings, key variables were aggregated into separate spreadsheets for analyses. The eye tracking variables that were calculated were dwell time, fixation count, and fixation time.

Dwell time was defined as the sum of all fixations and saccades. Because dwell time includes saccades, which could potentially inflate the results, fixation time and fixation count were also selected. Fixation time was defined as the sum of the fixation durations. Fixation count was defined as the number of all fixations. These three were selected to test whether the audio information in the matching condition influenced gaze towards that AOI. When testing the degree to which the audio statement matching one AOI (object) attracted visual attention, I compared the critical AOI to a mean of the other three non-critical AOI's. Gaze was quantified using all three measures – dwell time, fixation time, and fixation count.

## Results

Unless otherwise stated, all statistical analyses were conducted using SPSS (IBM Corp., 2013). I performed Shapiro-Wilk Tests for normality on variables of interest at the interval level. The results indicated that the DCCS, Day-Night Task and Stroop (all three variables measuring executive function for both groups) had negative skews which could not be fixed with transformations. Hence, I chose non-parametric tests for these variables. Additionally, the eye tracking variables (dwell time, fixation count, and fixation time) showed skew, but it was possible to transform them into more normal distributions using reverse-log transformations. Dwell time, fixation count, and fixation time were converted into proportion dwells, proportion fixations, and proportion fixation counts, so the outcomes do not represent the milliseconds (dwell time and fixation time) or fixation count directly, but a proportion value thereof.

Levene's Test for homogeneity of variance indicated that variances in dwell time with respect to critical AOI's were similar when comparing children, adolescents and adults, $F_{(2, 38)}$ = 3.47, *p* = .759. This was also true for fixation count on critical AOI's ($F_{(2, 38)}$ = 1.81, *p* = .512), fixation time on critical AOI's ($F_{(2, 38)}$ = 2.57, *p* = .970) and memory performances ($F_{(2, 38)}$ = 1.10, *p* = .523). The means, standard deviations, minimum and maximum scores by age group, including psychometric and experimental variables is shown in Table 1. It shows that the variables for executive function (DCCS, Day-Night, and Stroop) had roof-effects, and especially so with the DCCS.

*Table 1. n*, Mean, *SD*, Minimum and Maximum Scores for Study Variables by Age Group

| | Children *n* = 10 | | | Adolescents *n* = 15 | | | Adults *n* = 16 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | *SD* | Min/Max | Mean | *SD* | Min/Max | Mean | *SD* | Min/Max |
| Age | 6.4 | 1.6 | 4.1/10 | 15.2 | 1.0 | 14.9/17 | 23.9 | 2.1 | 21/28 |
| DCCS | 11 | 1.41 | 8/12 | - | - | - | - | - | - |
| Day-Night | 14.6 | 2.94 | 6/16 | - | - | - | - | - | - |
| Vocab | 25.4 | 3.72 | 17/30 | 23.6 | 5.78 | 11/33 | 39 | 5.34 | 25/50 |
| Info. | 22.7 | 5.37 | 11/29 | 15.33 | 3.40 | 11/20 | 20.4 | 2.5 | 16/25 |
| Stroop RT | - | - | - | 1.1 | .33 | .7/1.9 | 1.0 | .23 | .6/1.4 |
| Stroop Accuracy | - | - | - | 28.6 | 2.9 | 18/30 | 27 | 5.2 | 15/30 |
| Mem. Perform. | 18.30 | 3.04 | 14/24 | 14.53 | 3.6 | 8/22 | 15.4 | 3.7 | 9/22 |
| Dwell Time | 3.09 | .47 | 1.75/4.30 | 3.20 | .46 | 1.70/4.54 | 3.20 | .46 | 1.73/4.33 |
| Fix. Count | .77 | .41 | .0/1.90 | .84 | .39 | .0/2.3 | .85 | .41 | .0/2 |
| Fix. Time | 3.0 | .48 | 1.72/4.25 | 3.10 | .47 | 1.70/4.46 | 3.10 | .46 | 1.73/4.21 |

*Correlations between variables:* To see how the variables in the study related to one another, I calculated Spearman's Rho between all variables, for both the children and adolescent/adult groups. Spearman's Rho was selected because, as previously mentioned, some variables were not normally distributed. Tables 2 and 3 show the correlation coefficients.

For the children's group, correlations between age and psychometric variables were in the predicted directions, and the correlations with language variables (vocabulary and information), dwell time and fixation time were significant. Also, vocabulary had high correlations to information, memory performance, dwell time and fixation time while information had a high correlation to the DCCS measure of executive function but neither memory performance nor dwell and fixation time. In addition, the two measures of executive function (DCCS and Day-Night Task) were highly correlated with each other and the Day-Night Task was also correlated to memory performance, dwell time, fixation count, and fixation time. Dwell time was significantly correlated to both fixation count and fixation time. Finally, fixation count was significantly correlated with fixation time.

*Table 2*. Spearman's Correlation Coefficient (Rho) Between Variables for Children ($N$ = 10).

| | Age | Vocab. | Info. | DCCS | Day-Night | Memory Perform. | Dwell Time | Fix. Count | Fix. Time |
|---|---|---|---|---|---|---|---|---|---|
| Age | - | | | | | | | | |
| Vocab. | .79** | - | | | | | | | |
| Info. | .67* | .77** | - | | | | | | |
| DCCS | .51 | .63 | .84** | - | | | | | |
| Day-Night | .59 | .61 | .61 | .70* | - | | | | |
| Memory Perform. | -.094 | .73* | .54 | .40 | .65* | - | | | |
| Dwell Time | .34* | .67* | .59 | .57 | .86** | .14 | - | | |
| Fix. Count | .24 | .58 | .51 | .52 | .78** | -.06 | .81** | - | |
| Fix. Time | .34* | .65* | .55 | .55 | .84** | .16 | .73** | .73** | - |

*p<.05, **p<.01

For the adolescent/adult group, correlations between age and psychometric variables were largely expected, but only the associations with language variables (vocabulary and information) and dwell time were significant. Additionally, vocabulary was highly correlated to information and weakly correlated to memory performance. Finally, dwell time was highly correlated to fixation count and fixation time while fixation count was correlated with fixation time.

*Table 3*. Spearman's Correlation Coefficient (Rho) Between Variables for Adolescents and Adults (*N* = 31).

| | Age | Vocab. | Info. | Stroop Accuracy | Stroop RT | Memory Perform. | Dwell Time | Fix. Count | Fix. Time |
|---|---|---|---|---|---|---|---|---|---|
| Age | - | | | | | | | | |
| Vocab. | .79** | - | | | | | | | |
| Info. | .55** | .81** | - | | | | | | |
| Stroop Accuracy | .05 | -.05 | .03 | - | | | | | |
| Stroop RT | .04 | -.03 | -.30 | -.04 | - | | | | |
| Memory Perform. | -.09 | .39* | .30 | .14 | -.23 | - | | | |
| Dwell Time | .34* | .05 | .20 | .03 | -.19 | .137 | | | |
| Fix. Count | .24 | .04 | .08 | -.11 | .15 | -.06 | .81** | | |
| Fix. Time | .34* | .05 | .23 | .02 | -.28 | .16 | .97** | .73** | - |

*p<.05, **p<.01

## Condition and Age on Memory Performance

I conducted a one-way repeated measures analysis of variance to compare the effects of experimental condition (control, matching, non-matching) within-subjects and age-group (children, adolescent, and adult) between-subjects on memory performance. Mauchly's test of sphericity, which tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variable is proportional to an identity matrix, indicated a violation of the assumption of sphericity. The Greenhouse-Geisser criterion was applied when the test indicated a violation of the assumptions of sphericity. There was no significant effect of condition on memory performance, Wilk's Lambda = .916, $F_{(2, 71)} = 1.36$, $p = .262$. While not significant, the memory performances within-subjects by condition were in the expected directions, with the lowest performing condition being the non-matching third condition (*M*=1.267, *SE*= .066), the highest performing condition being the matching second condition (*M*=1.396, *SE*=.066) and the control first condition falling between the two (*M*=1.355, *SE*=.069). Furthermore, there was no interaction effect of condition and age group, $F_{(4, 72)} = .518$, $p = .712$.

There was a statistically significant difference of age group on memory performance, $F_{(2, 38)} = 3.361$, $p = .045$. Post-hoc comparisons using the LSD test indicated that the children had significantly higher memory performance (*M* = 1.53, *SD* = .27) than the adolescent group (*M* = 1.21, *SD* = .31) at the .05 level. However, neither the memory performance of children nor of adolescents was significantly different from that of adults (*M* = 1.28, *SD* = .32).


*Figure 3.* Memory Performance Means for children, adolescent, and adult groups, by Condition.

**Mediation of Executive Function**

I calculated a multiple linear regression to predict memory performance in the distraction condition based on executive function variables of the DCCS and Day-Night Task for the children's group. The EF variables captured over a third of variance in memory performance ($R^2$ of .373) but the model was not significant, $F_{(2, 7)} = 2.08$, $p = .196$. Since the EF variables were not normally distributed, the indirect effect was tested using a bootstrap estimation approach with 1000 samples (Shrout & Bolger, 2002). These results indicated the indirect coefficient was not significant for the DCCS, b = .086 *SE* = .056, 95% CI=-.047, .218 and the Day-Night Task, b = .013 *SE* =.017, 95% CI = -.027, .053.

I performed the same analysis to predict memory performance in the distraction condition based on executive function variables of the Stroop Accuracy and Stroop Response Time for the adolescent and adult age group. The model was not significant, $F_{(2, 60)} = 1.73$, $p = .186$ with an $R^2$ of .018. Bootstrap estimation approach results indicated also non-significant indirect coefficients for the Stroop, b = .013 *SE* = .011, 95% CI =-.010, .036 and Stroop Response Time, b=-.298 SE=.207, 95% CI=-.712, .116.

**Audio-visual match effects on eye tracking variables**

I conducted a paired samples t-test to compare dwell time, fixation count, and fixation time on AOI's in which the audio information and image matched and on AOI's in which the audio information and image did not match during encoding. There was no significant difference in the dwell time of audio information/image-non-matched AOI's (*M* = 3.16, *SD* = .17), and audio information/image-matched AOI's (*M* = 3.14, *SD* = .18), $t_{(39)} = .636$, $p = .529$. There was not a significant difference in the fixation count of audio information/image-non-matched AOI's (*M* = .80, *SD* = .16), and audio information/image-matched AOI's (*M* = .80, *SD* = .16), $t_{(39)} = .35$, $p = .732$. Finally, there was not a significant difference in the fixation time of audio information/image-non-matched AOI's (*M* = 3.10, *SD* = .20), and audio information/image-matched AOI's (*M* = 3.03, *SD* = .23), $t_{(39)} = 1.11$, $p = .273$.

I conducted a one-way repeated measures analysis of variance to compare the effects of audio information/image match to AOI and audio information/image non-match to AOI within-subjects and age-group (children, adolescent, and adult) between-subjects on dwell time of matched and not-matched AOI's during encoding. Mauchly's test of sphericity indicated a violation of the assumption of sphericity, so the Greenhouse-Geisser criterion was applied. There was not a statistically significant effect, Wilk's Lambda = .983, $F_{(1, 38)} = .65$, $p = .427$.

While not significant, the dwell times within-subjects on audio information/image matched AOI's and audio information/image non-matched AOI's were expected, with the longer dwells being on the non-matched AOI's ($M = 3.15$, $SE = .026$), and the shorter dwells being on the matched AOI's ($M = 3.13$, $SE = .027$). Furthermore, there was no interaction effect of match and age group, $F_{(2, 38)} = .651$, $p = .527$.

There was a statistically significant difference of age group on dwell time of audio information/image matching AOI's and audio information/image-non-matching AOI's, $F_{(2, 38)} = 4.42$, $p = .019$. Post-hoc comparisons using the LSD test indicated that the children had significantly shorter dwell times ($M = 3.04$, $SD = .04$) than the adolescent group ($M = 3.20$, $SD = .04$) and the adult group ($M = 3.20$, $SD = .04$) at the .05 level. However, the dwell times of adolescents were not significantly different from that of adults.

I conducted a one-way repeated measures analysis of variance to compare the effects of audio information/image match to AOI and audio information/image non-match to AOI within-subjects and age-group (children, adolescent, and adult) between-subjects on fixation count of matched and not-matched AOI's during encoding. Mauchly's test of sphericity indicated a violation of the assumption of sphericity, so the Greenhouse-Geisser criterion was applied. There was not a statistically significant effect, Wilk's Lambda = .994, $F_{(1, 38)} = .21$, $p = .647$.

While not significant, the fixation counts within-subjects on audio information/image matched AOI's and audio information/image non-matched AOI's were expected, with more fixations being on the non-matched AOI's (*M* = .80, *SE* = .026), and fewer fixations being on the matched AOI's (*M* = .79, *SE* = .024). Furthermore, there was no interaction effect of match and age group, $F_{(2, 38)}$ = .326, *p* = .724.

There was not a statistically significant difference of age group on fixation count of audio information/image matching AOI's and audio information/image-non-matching AOI's, $F_{(2, 38)}$ = 2.87, *p* = .069. While not statistically significant, it was approaching significance. Post-hoc comparisons using the LSD test indicated that the children had fewer fixations (*M* = .71, *SD* = .04) than the adolescent group (*M* = .83, *SD* = .04) and the adult group (*M* = .83, *SD* = .04).

Finally, I conducted a one-way repeated measures analysis of variance to compare the effects of audio information/image match to AOI and audio information/image non-match to AOI within-subjects and age-group (children, adolescent, and adult) between-subjects on fixation time of matched and not-matched AOI's during encoding. Mauchly's test of sphericity indicated a violation of the assumption of sphericity, so the Greenhouse-Geisser criterion was applied. There was not a statistically significant effect, Wilk's Lambda = .945, $F_{(1, 38)}$ = 2.13, *p* = .153.

While not significant, the fixation times within-subjects on audio information/image matched AOI's and audio information/image non-matched AOI's were expected, with the longer fixations being on the non-matched AOI's (*M* = 3.05, *SE* = .032), and the shorter dwells being on the matched AOI's (*M* = 3.01, *SE* = .034). Furthermore, there was no interaction effect of match and age group, $F_{(2, 38)}$ = 1.74, *p* = .189.

There was a statistically significant difference of age group on fixation time of audio information/image matching AOI's and audio information/image-non-matching AOI's, $F_{(2, 38)}$ = 3.5, *p* = .041. Post-hoc comparisons using the LSD test indicated that the children had significantly shorter fixation times (*M* = 2.91, *SD* = .06) than the adolescent group (*M* = 3.10, *SD* = .05) and the adult group (*M* = 3.10, *SD* = .05) at the .05 level. However, the fixation times of adolescents were not significantly different from that of adults.

## Discussion

The current study investigated how the presence of relevant and irrelevant spoken language affects the processing for visual information, and if this is mediated by executive function processes that develop from childhood to adolescence. It was hypothesized that when participants were listening to relevant spoken language, their memory performance would be significantly enhanced. It was also hypothesized that when participants were listening to irrelevant spoken language, their memory performance would be lowered, and that memory performance during trials with irrelevant spoken language would be mediated by executive function. The results did not support these hypotheses. Furthermore, it was hypothesized that age would significantly impact memory performance, whereby adolescents and adults would perform similarly and better than the children. Unexpectedly however, while there was a significant impact of age on memory performance, children performed the best, followed by adults and adolescents.

Furthermore, I hypothesized that audio/visual information matched to AOI would lead to shorter dwell times, fixation times, and fewer fixations on the matching AOI while audio/visual information non-matched to AOI would lead to longer dwell times, fixation times, and more fixations on the non-matching AOI's. This was not supported. Finally, it was hypothesized that age group would impact dwell time, fixation count, and fixation time, such that children would have shorter dwells and fixations as well as fewer fixations. Results supported this hypothesis.

**Differences within Conditions and between Age on Memory Performance**

While memory performances by condition were in the expected directions, with the lowest performing condition being the non-matching third condition, the highest performing condition being the matching second condition, and the control first condition falling between the two, these differences were not significant, as had been hypothesized. In addition, memory performances by age-group were somewhat unexpected, as the highest performing group was the children, followed by the adults and then the adolescents. This was contrary to predictions that the lowest performing group would be the children, while the adolescent and adult groups would perform similarly.

There is a plethora of probable reasons as to why there was no significant difference of condition on performance in the memory task while there was a significant difference of age group on performance, with the children performing the best. As mentioned previously, I decided to add the cues for the children's group, but not the adolescent or adult group. This decision resulted in it being nearly impossible to compare the results on the memory performance of the child age group to the adolescent and adult groups. In short, the integrity of the memory performance scores was called into question. However, at the time of pilot testing, it seemed the only reasonable measure to take to keep the children feeling motivated throughout the duration of the experiment and able to complete the full experiment. While it was initially intended as a last resort to offer the cues if a child participant was struggling repeatedly and had given three consecutive incorrect answers, it then seemed unfair not to continue with the practice for all children participants equally. Adding such hints seemed to help the children participants feel more engaged during the experiment, as they more frequently got correct answers with the cues available to them. Additionally, I decided to break up the main memory task into two parts, as it was overly challenging to ask the children to focus on the full task at one time. Because of these alterations to the experimental procedure, no children failed to complete all 12 trials of the experiment, which was great. However, it lowered the integrity of the performance measure for the children age group.

Moreover, since the condition differences on memory performance were in the expected order, but not of significance, it is likely that the effect was too small. Another contributor could be that there were too few trials in each condition – only four. Given that the instructions were to look at and remember the pictures, with no mention of the audio aspect, it is possible that participants quickly realized that listening to the audio was not the primary task and ignored it altogether, actively suppressing it as irrelevant stimuli. Research by Friedrich and colleagues (2011) showed that users of a brain-computer interface could maintain control during an imagery task while ignoring tones in a passive distraction situation (completely ignore auditory stimuli), increasing their performance as compared with an active distraction condition (react to target stimuli by pressing a button). In a study comparing the impact of auditory load on visual distractors, and of visual load on auditory distractors within the same individuals, it was found that auditory processing was more impacted by visual distraction than the other way around (Tellinghuisen, & Nowak, 2003). Overall, it seems that the audio did not adequately serve as active distraction or active facilitation.

Instead, it is likely that the audio served as a passive distraction or background noise. A better test could have been where the distraction was actively competing with information that had to be ignored. Research has shown that humans are especially susceptible to attend to stimuli of biological significance, such as threatening stimuli (Öhman, Flykt, & Esteves, 2001). Hence, it could be interesting to investigate whether such stimuli can capture attention beyond top-down control.

Initially, I had planned to make the task more challenging for the matching and non-matching conditions by asking participants to recall the sentence which they had heard.

Adding this extra element would have made the act of listening to the audio more salient and taxing cognitively. While this addition would have likely been feasible for the adolescent and adult groups, it was deemed overly demanding on the children's group and removed from the design. Another consideration was to vary what should have been recalled, such that on some trials you had to recall the spoken sentence while on others it was the images that needed to be recalled. However, I decided to go with a simpler design in which the visual images were task-relevant and the audio was related to or unrelated to objects for later recall. Overall, it was quite difficult to find a balance between an experiment that would be feasible for all age-groups while also being challenging enough, but not too challenging, for the various groups.

**Mediation of Executive Function**

When participants were in the non-matching third condition, I hypothesized that memory performance would be mediated by executive function. However, this was not supported. Executive function scores were not a reliable predictor of memory performance scores. This could be due to problems with the tests. First, DCCS showed strong roof-effects in the present study, which undermines its credibility as a measure of executive function. As Wiebe, Morton, Buss, and Spencer (2014) have shown, although 5-year-olds have little trouble switching rules, most 3-year-olds (typically around 70%) perseverates and continues using the first set of rules after they are instructed to switch. However, the children tested in the present study were rather few ($N$=10) and only one child was below the age of five. As a result, most children performed the DCCS with ease, giving little variability in scores. Second, while the Day-Night task had slightly more variability in scores, it was not able to encapsulate response time scores. When testing with the Day-Night task, many children could produce the correct result, but only because they spent a large amount of time actively suppressing the name of the image and then saying the opposite. The number of correct answers which was recorded only tells part of the story of their EF abilities, because it does not reflect the speed with which they were able to suppress the incorrect answers. In addition, critics have pointed out that features of the Day-Night task place extraneous demands on children's memory, thereby weakening its measurement of inhibitory control (Vendetti, Kamawar, Podjarny, & Astle, 2015).

Furthermore, the mediation of executive function was not supported for the adolescent and adult groups. Since it seems likely that the distracting condition was not distracting enough, then it seems clear that that there would be no need for EF to control attention. Hence, within the context of the lack of a main effect of condition on memory performance as seen in the previous section, the lack of mediation of executive function scores can be said to be expected.

**Audio-visual match effects on eye tracking variables**

While dwell times and fixation times were shorter, and fixation counts fewer in the case of a match between audio information and image on AOI as compared with those on a non-match between audio information and image on AOI, these differences were not statistically significant. There are several possible explanations for why there was no significant difference of a match between audio and visual information on eye tracking variables. This seems to confirm previous research showing biasing of attention away from previously attended stimuli and toward new ones (Thomas & Llena, 2009). However, previous work has focused on the biasing of attention away from visual stimuli previously encountered and toward new visual stimuli while the present study seemed to show a biasing of attention away from auditory stimuli. Alternatively, as discussed earlier, it is possible that participants actively ignored all audio information (both facilitating and distracting) and focused exclusively on the visual stimuli to be remembered.

Finally, it was shown that age group significantly impacted dwell times, fixation counts, and fixation times on both matching and non-matching AOI's, such that children had shorter dwells and fixations as well as fewer fixations overall as compared with both adolescents and adults. Adolescents and adults had similar eye tracking outcomes for dwell times, fixation times and fixation counts. This shows support for previous research showing as children get older they get better at sustaining attention for longer intervals and become less susceptible to interference from irrelevant distractors (Matusz et al., 2015), although attentional abilities vary between children of the same age.

However, it is interesting to note that the dwell times, fixation times and fixation counts of adolescents and adults were nearly identical, supporting the idea that adolescents are equally capable (or incapable) as adults of sustaining attention for longer periods of time. Developmental research into adolescents shows that this transition period from childhood to adulthood involves greater use of cognitive skills, such as the use of top-down effortful control to modify attention. Longitudinal research examining changes in brain structure over development within individuals has shown that cortical white matter increases roughly linearly with age throughout childhood and adolescence, and differs little across regions (Gogtay et al., 2004; Sowell et al., 2004). Conversely, cortical grey matter, which mirrors neuronal density and the number of connections between neurons, follows an inverted-U shape over development, peaking at different ages depending on the region. Therefore, grey matter loss is considered an index of the time-course of maturation (Sowell et al., 2004). More recent work has shown distinct patterns of development of cortical volume, thickness and surface area during adolescence. This includes a combination of linear and quadratic trajectories of change in the cortex. Specifically, surface area showed non-linear increases across adolescence over most of the cortical area while thickness was characterized by both linear and non-linear decreasing trajectories across much of the frontal and parietal regions, as well as non-linear increases in temporo-occipital regions. On the other hand, volume showed non-linear increases in parts of the frontal and temporo-occipital regions, as well as linear reductions within frontal and parietal regions. (Vijayakumar, et al., 2016). Such findings show a hierarchical pattern of maturation of brain development, supporting the gradual emergence of adult-like inhibitory control that the present findings highlight.

**Directions for Future Research**

The results highlight the need for more research in which there is more active competition between goal-relevant information and goal-irrelevant information in order to see the extent to which distractor processing can be prevented. As stated previously, it could be interesting to investigate whether biologically salient stimuli, such as threatening stimuli, can capture attention beyond top-down control. Additionally, further research should investigate more direct links to how and when the verbal information should affect gaze. It is also important that future studies experiment with other possibly related aspects. For instance, it would be interesting to compare susceptibility to distraction of adults with spoken sentences in a foreign language versus a native language.

**Acknowledgements**

The author would like to thank Elia Psouni and Roger Johansson for their valuable insight and feedback throughout the duration of the project. I would like to thank Rachel Elsinga for her contributions to this project. Also, I would like to thank the Humanities Laboratory at Lund University for allowing me to borrow their eye tracking equipment and keep it at the Psychology Department over the duration of the project. Finally, I would like to thank all the participants (and their caregivers) for taking their valuable time to participate in my study.

**References**

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, *38*(4), 419-439.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *The Quarterly journal of experimental psychology*, *24*(2), 225-235.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 247–264.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518. http://doi.org/10.1016/j.jml.2006.12.004

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology*, *49*(1), 5–28.

Baddeley, A., & Hitch, G. J. (1994). Developments in the Concept of Working Memory. *Neuropsychology*, *8*(4), 485–493.

Barkley, R. A. (1998). Attention deficit hyperactivity disorder. A handbook for diagnosis and treatment. New York: Guilford

Barnett, V., & Lewis, T. (1978). *Outliers in statistical data.* New York: Wiley.

Blair, C., & Razza, R. P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development*, 647–663.

Braddick, O., Atkinson, J., & Innocenti, G. (2011). Neurodevelopment of the visual system in typically developing children. *Gene Expression to Neurobiology and Behaviour: Human Brain Development and Developmental Disorders*, *189*, 113.

Broadbent, D. E. (1958). Perception and communication. London: Pergamon Press.

Campbell, S. B., & Von Stauffenberg, C. (2009). Delay and inhibition as early predictors of ADHD symptoms in third grade. *Journal of Abnormal Child Psychology*, *37*(1), 1-15.

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development*, *22*(4), 489–510. http://doi.org/10.1016/j.cogdev.2007.08.002

Carlson, S. M., Faja, S., & Beck, D. M. (2015). Incorporating early development into the measurement of executive function: The need for a continuum of measures across development. In *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research.* (pp. 45–64). Washington: American Psychological Association. http://doi.org/10.1037/14797-003

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive Function and Theory of Mind: Stability and Prediction from Ages 2 to 3. *Developmental Psychology*, *40*(6), 1105–1122. http://doi.org/10.1037/0012-1649.40.6.1105

Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., et al. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences*, *108*(36), 14998–15003. http://doi.org/10.1073/pnas.1108561108

Casey, B. J., Tottenham, N., & Fossella, J. (2002). Clinical, imaging, lesion, and genetic approaches toward a model of cognitive control. *Developmental Psychobiology*, *40*(3), 237–254. http://doi.org/10.1002/dev.10030

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, *37*(5), 715–730. http://doi.org/10.1037//0012-1649.37.5.715

Cooper, R. M. (1974). Cooper, 1974. *Cognitive Psychology*, 84–107.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review*, *70*(1), 80.

Enns, J., & Brodeur, D. (1989). A developmental study of covert orienting to peripheral visual cues. Journal of Experimental Child Psychology, 48, 171–189. doi:10.1016/0022-0965(89)90001-5

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Attention, Perception, & Psychophysics*, *16*(1), 143-149.

Erickson, L. C., Thiessen, E. D., Godwin, K. E., Dickerson, J. P., & Fisher, A. V. (2015). Endogenously and exogenously driven selective sustained attention: Contributions to learning in kindergarten children. *Journal of experimental child psychology*, *138*, 126-134.

Friedrich, E. V., Scherer, R., Sonnleitner, K., & Neuper, C. (2011). Impact of auditory distraction on user performance in a brain–computer interface driven by different mental tasks. *Clinical Neurophysiology*, *122*(10), 2003-2009.

Gaspelin, N., Margett-Jordan, T., & Ruthruff, E. (2015). Susceptible to distraction: Children lack top-down control over spatial attention capture. *Psychonomic Bulletin & Review*, *22*(2), 461–468. http://doi.org/10.3758/s13423-014-0708-0

Gatti SV, Egeth HE (1978). Failure of spatial selectivity in vision. Bull Psychon Soc 11:181–184

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 312–7 years old on a Stroop-like day-night test. *Cognition*, *53*(2), 129-153.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., ... & Rapoport, J. L. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National academy of Sciences of the United States of America*, *101*(21), 8174-8179.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. *Eye movements: A window on mind and brain*, 537-562.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-B32.

Hughes, C., & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, *108*(3), 663–676. http://doi.org/10.1016/j.jecp.2010.06.005

IBM Corp. (2013). IBM SPSS Statistics for Mac, Version 22.0. Armonk, NY: IBM Corp.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. *Varieties of attention*, *1*, 29-61.

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child development*, *68*(2), 263-277.

Lahey, B. B., Lee, S. S., Sibley, M. H., Applegate, B., Molina, B. S., & Pelham, W. E. (2016). Predictors of adolescent outcomes among 4–6-year-old children with attention-deficit/hyperactivity disorder. *Journal of abnormal psychology*, *125*(2), 168.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in*

*Cognitive Sciences*, *9*(2), 75–82. http://doi.org/10.1016/j.tics.2004.12.004

Malpass, D., & Meyer, A. S. (2010). The time course of name retrieval during multiple-object naming: Evidence from extrafoveal-on-foveal effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 523.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition*, *136*, 156-165.

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & cognition*, *37*(7), 1026-1039.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693–2698. http://doi.org/10.1073/pnas.1010076108/-/DCSupplemental/sapp.pdf

Mordre, M., Groholt, B., Sandstad, B., & Myhre, A. M. (2012). The impact of ADHD symptoms and global impairment in childhood on working disability in mid-adulthood: a 28-year follow-up study using official disability pension records in a high-risk in-patient population. *BMC psychiatry*, *12*(1), 174.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological review*, *75*(6), 522.

Pardini, D. A., & Fite, P. J. (2010). Symptoms of conduct disorder, oppositional defiant disorder, attention-deficit/hyperactivity disorder, and callous-unemotional traits as unique predictors of psychosocial maladjustment in boys: advancing an evidence base for DSM-V. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(11), 1134-1144.

Rayner, K. (2009). *Eye movements and attention in reading, scene perception, and visual search*. *Q J Exp Psychol (Hove)* (Vol. 62). http://doi.org/10.1080/17470210902816461

Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition*, *33*(8), 1325-1336.

Schmitter-Edgecombe, M. (1996). The effects of divided attention on implicit and explicit memory performance. *Journal of the International Neuropsychological Society*, *2*(2), 111-125.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting Adolescent Cognitive and Self-Regulatory Competencies From Preschool Delay of Gratification: Identifying Diagnostic Conditions. *Developmental Psychology*, 978–986.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological methods*, *7*(4), 422-445.

Skodzik, T., Holling, H., & Pedersen, A. (2017). Long-term memory performance in adult ADHD: a meta-analysis. *Journal of attention disorders*, *21*(4), 267-283.

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.

Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW (2004): Longitudinal mapping of cortical thickness and brain growth in normal children. J Neurosci 24:8223–8231.

Spira, E. G., and Fischel, J. E. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: a review. *J. Child Psychol. Psychiatry* 46, 755–773. doi: 10.1111/j.1469-7610.2005.01466.x

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Tanenhaus et al., 1995. *Science*, 1632–1634.

Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Attention, Perception, & Psychophysics*, *65*(5), 817-828.

Thomas, L. E., & Lleras, A. (2009). Inhibitory tagging in an interrupted visual search. *Attention, Perception, & Psychophysics*, *71*(6), 1241-1250.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Ungerleider, L., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549-586). Cambridge, MA: The MIT Press.

Vendetti, C., Kamawar, D., Podjarny, G., & Astle, A. (2015). Measuring Preschoolers' Inhibitory Control Using the Black/White Stroop. *Infant and Child Development, 24*(6), 587-605. doi: 10.1002/icd.1902

Vijayakumar, N., Allen, N. B., Youssef, G., Dennison, M., Yücel, M., Simmons, J. G., & Whittle, S. (2016). Brain development during adolescence: A mixed-longitudinal investigation of cortical thickness, surface area, and volume. *Human brain mapping*, *37*(6), 2027-2038.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*, *22*, 498.

Wechsler, D. (2012). Wechsler preschool and primary scale of intelligence—fourth edition (WPPSI-IV).

Wiebe, S. A., Morton, J. B., Buss, A. T., & Spencer, J. P. (2014). The emergent executive: A dynamic field theory of the development of executive function: II. A case study of the DCCS task. *Monographs of the Society for Research in Child Development*, *79*(2), 12-25.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*, *57*(11), 1336-1346.

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of Inhibitory Control Across the Life Span. *Developmental Psychology*, 205–213.

Yarbus, A. L. (1967). Eye movements and vision. 1967. *New York*.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature protocols*, *1*(1), 297.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of experimental psychology: general*, *130*(3), 466.