University of Birmingham

Crossmodal spatial distraction across the lifespan

Pedale, Tiziana; Mastroberardino, Serena; Capurso, Michele; Bremner, Andrew J.; Spence, Charles; Santangelo, Valerio

[Link to publication on Research at Birmingham portal](#)

Running head: CROSSMODAL DISTRACTION ACROSS THE LIFESPAN

Crossmodal spatial distraction across the lifespan

Tiziana Pedale[1], Serena Mastroberardino[2], Michele Capurso[3],

Andrew J. Bremner[4], Charles Spence[5], & Valerio Santangelo[1,3]

1. Neuroimaging Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy
2. Department of Psychology, Sapienza University of Rome, Italy
3. Department of Philosophy, Social Sciences & Education, University of Perugia, Italy
4. School of Psychology, University of Birmingham, UK
5. Department of Experimental Psychology, Oxford University, UK

<u>Corresponding Author</u>:
Valerio Santangelo,
Dept. of Philosophy, Social Sciences & Education,
University of Perugia,
Piazza G. Ermini, 1,
06123 Perugia (Italy)
E-mail: valerio.santangelo@unipg.it

ABSTRACT

The ability to resist distracting stimuli whilst voluntarily focusing on a task is fundamental to our everyday cognitive functioning. Here, we investigated how this ability develops, and thereafter declines, across the lifespan using a single task/experiment. Young children (5-7 years), older children (10-11 years), young adults (20-27 years), and older adults (62-86 years) were presented with complex visual scenes. Endogenous (voluntary) attention was engaged by having the participants search for a visual target presented on either the left or right side of the display. The onset of the visual scenes was preceded – at stimulus onset asynchronies (SOAs) of 50, 200, or 500 ms – by a task-irrelevant sound (an exogenous crossmodal spatial distractor) delivered either on the same or opposite side as the visual target, or simultaneously on both sides (cued, uncued, or neutral trials, respectively). Age-related differences were revealed, especially in the extreme age-groups, which showed a greater impact of crossmodal spatial distractors. Young children were highly susceptible to exogenous spatial distraction at the shortest SOA (50 ms), whereas older adults were distracted at all SOAs, showing significant exogenous capture effects during the visual search task. By contrast, older children and young adults' search performance was not significantly affected by crossmodal spatial distraction. Overall, these findings present a detailed picture of the developmental trajectory of endogenous resistance to crossmodal spatial distraction from childhood to old age and demonstrate a different efficiency in coping with distraction across the four age-groups studied.

Keywords: Visual search, complex scenes, crossmodal spatial attention, stimulus onset asynchrony, cognitive development.

1. INTRODUCTION

In our everyday life, we are surrounded by continuously changing streams of information that compete to access our attention through different sensory channels (Bundesen, Habekost, & Kyllingsbæk, 2011). Attentional selection is often described as the result of the operation of two qualitatively different attentional mechanisms, one reflexive or stimulus-driven (exogenous attention) and the other voluntary or goal-driven (endogenous attention) (Buschman & Miller, 2007; Desimone & Duncan, 1995; Spence & Driver, 2004; Theeuwes, 2010). In this article, we describe an innovative lifespan investigation designed to address the development of the interactive roles which exogenous and endogenous attention play in determining our distractibility during attentional selection.

Research shows that attentional resources are often deployed in a way that reflects an interactive relationship between mechanisms of exogenous and endogenous attention. For instance, several studies have demonstrated that exogenous attention can be modulated by endogenous factors related to expectations and task demands (e.g., Santangelo, Olivetti Belardinelli, & Spence, 2007; Santangelo, Olivetti Belardinelli, Spence, & Macaluso, 2009; Spalek, Falcon, & Di Lollo, 2006; Yantis & Jonides, 1990; for reviews, see Macaluso & Doricchi, 2013; Santangelo & Macaluso, 2012; Santangelo & Spence, 2008; see, however, Theeuwes, 2010; Van der Lubbe & Postma, 2005, for evidence suggesting that exogenous attention to salient stimuli is entirely reflexive and automatic). The interplay between endogenous and exogenous attention occurs throughout our daily lives. During most of our daily activities, our senses are under constant bombardment by a wide array of different sources of information, many of which are typically irrelevant to the task at hand. As a consequence, endogenous attention devoted to currently relevant information is frequently challenged by distracting stimuli that might capture available attentional resources exogenously and divert them from our goals. On the other hand, increased endogenous attention to task demands can reduce – under at least certain circumstances – the exogenous capture that is driven by task-irrelevant distractors (Forster & Lavie, 2009; Lavie, Hirst, De Fockert, & Viding, 2004; see also Santangelo et al., 2007; Yantis & Jonides, 1990).

Task-relevant and task-irrelevant stimuli may be presented in the same or different sensory modalities. For example, the presentation of a task-irrelevant sound from a particular spatial location may exogenously attract attention to that region of space, thereby: (i) facilitating the processing of relevant visual stimuli appearing in the same area, and also (ii) interfering with the selection of relevant visual stimuli located elsewhere (this phenomenon is known as crossmodal exogenous distraction; see Driver & Spence, 1998; McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003; McDonald & Ward, 2000; Spence & Driver, 2004; Spence & Soto-Faraco, 2020; Störmer, 2019; Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015). Indeed, even task-irrelevant sounds have been shown to activate the contralateral visual cortex, by eliciting an Auditory-evoked Contralateral Occipital Positivity (ACOP; Retsa, Matusz, Schnupp, & Murray, 2020), in line with the notion that visuospatial exogenous attention is enhanced at the location of the auditory stimulus (Hillyard, Störmer, Feng, Martinez, & McDonald, 2016). A critical factor determining whether exogenous stimuli have facilitatory and/or interfering effects is the temporal interval between the appearance of the task-irrelevant stimulus (i.e., the exogenous "cue"), and the task-relevant target. This is known as the "stimulus onset asynchrony" (SOA). At short SOAs (ranging between 50 and 300 ms), exogenous cues tend to increase their ability to capture attention spatially, resulting in the enhanced detection of relevant stimuli appearing in the same area (Berger, Henik, & Rafal, 2005; McDonald & Ward, 1999, 2000; Spence & McDonald, 2004; Van der Stoep et al., 2015). But at SOAs in excess of 300 ms, exogenous benefits typically tend to become smaller, enabling the so-called "inhibition of return" (IOR) effect (Klein, 2000; Klein & MacInnes, 1999; Lupiáñez, Klein, & Bartolomeo, 2006; Spence, Lloyd, McGlone, Nicholls, & Driver, 2000). This latter effect is thought to be due to an initial reflexive shift of exogenous spatial attention to the cued location, followed by subsequent suppression in the processing of information from that location, with the consequent reduction in the facilitation for targets appearing there (Posner & Cohen, 1984). Both exogenous facilitation and IOR occur independently of sensory modality (e.g., Pierce, McDonald, & Green, 2018; Spence & Driver, 1998;

Spence et al., 2000; Tassinari & Campara, 1996) and appear to be preserved across age (Langley, Friesen, Saville, & Ciernia, 2011; MacPherson, Klein, & Moore, 2003).

Although exogenous attention has been shown to be relatively stable across development (Iarocci, Enns, Randolph, & Burack, 2009; Langley et al., 2011; Trick & Enns, 1998; Waszak, Li, & Hommel, 2010), endogenous attention to task-relevant stimuli has been shown to change according to a U-shaped pattern across the lifespan, with gradual improvement from childhood to adulthood, followed by a decline into old age (Enns, Brodeur, & Trick, 1998; Plude, Enns, & Brodeur, 1994; Rueda et al., 2004; Waszak et al., 2010). These developmental changes are thought to reflect a corresponding improvement and subsequent decline in the ability to endogenously inhibit task-irrelevant information from childhood to adulthood (e.g., Cavallina, Puccio, Capurso, Bremner, & Santangelo, 2018; Dempster, 1992; Harnishfeger, 1995) to old age (e.g., Dempster, 1992; Hasher & Zacks, 1988; Poliakoff, Ashworth, Lowe, & Spence, 2006; Tipper, 1991). However, exogenous attentional capture under concurrent endogenous attention control – or, adopting more ecological terminology, the resistance to distraction when we are voluntarily engaged in an attentionally-demanding task – has rarely been assessed across the lifespan. In fact, the majority of the literature in this area has addressed the development of exogenous and/or endogenous attention mechanisms in separate experiments, focusing on either one mechanism or the other (e.g., Folk & Hoyer, 1992; Iarocci et al., 2009; Leclercq & Siéroff, 2013). Furthermore, many studies have focused just on development in early life, or from adulthood to old age (e.g., Cavallina et al., 2018; Poliakoff et al., 2006; Rueda et al., 2004).

To date, only a few studies have used a single paradigm to investigate exogenous capture of spatial attention under concurrent endogenous control across the entire life span (e.g., Parker & Robinson, 2018; Robinson, Hawthorn, & Rahman, 2018). This literature appears to reveal that children, in comparison to adults, show reduced endogenous control in the presence of exogenous distraction (Cavallina et al., 2018; Gumenyuk, Korzyukov, Alho, Escera, & Näätänen, 2004; Matusz et al., 2015; Parker & Robinson, 2018; Robinson et al., 2018). In studies of aging in adulthood, the

available findings are not fully consistent. An account in terms of a higher susceptibility to auditory distraction in older adults (de Fockert, Ramchurn, Van Velzen, Bergström, & Bunce, 2009; Healey, Campbell, & Hasher, 2008) is supported by crossmodal oddball studies indicating that older adults have a visual discrimination impairment in the presence of a concurrent stream of rare and novel auditory distractors (Alain & Woods, 1999; Andrés, Parmentier, & Escera, 2006). However, other studies suggest an asymmetrical pattern of results depending on the sensory modality of the distracting stimulus. Equivalent magnitudes of auditory distraction in young and older adults were found when attending to task-relevant stimuli in the visual modality (Guerreiro, Murphy, & Van Gerven, 2013; Guerreiro & Van Gerven, 2011; Hugenschmidt, Peiffer, McCoy, Hayasaka, & Laurienti, 2009; Parker & Robinson, 2018; Robinson et al., 2018; Van Gerven & Murphy, 2010). By contrast, stronger distraction was revealed in older as compared to young adults in the presence of visual distractors when attending to auditory stimuli (Parker & Robinson, 2018; Van Gerven & Guerreiro, 2016), suggesting a visual dominance in aging during crossmodal stimulation (Barnhart, Rivera, & Robinson, 2018). Further research is therefore required to gain a clear picture of how the mechanisms regulating endogenous attention under concurrent exogenous distraction develop across the lifespan. Moreover, to the best of our knowledge, no studies have yet focused on the interplay between crossmodal spatial distraction and the control of endogenous attention across the entire lifespan – that is, from childhood to old age – in a single experiment. We report such an undertaking here with the aim of gleaning a more complete picture of the development and decline of endogenous resistance to crossmodal spatial distraction.

While the majority of the previous developmental literature in this area has used simple and repetitive stimuli, there are concerns that such abstract scenarios could create unequal task demands across distinct age groups. In order to foster as fair a comparison as possible, we therefore had our participants [young children, 5-7 years-old (y.o.); older children, 10-11 y.o.; young adults, 20-27 y.o., and older adults, 62-86 y.o.] perform a more ecological task (e.g., Felsen & Dan, 2005; Peelen & Kastner, 2014) which was designed to be more equally engaging across age groups, based on complex

and unrepeated visual scenes. Each scene included only one agentive element (i.e., an animal or a human being), which constituted the to-be-searched-for-target, located on either the right or left of the scene. The task was to localize the target (i.e., a voluntary – endogenous – spatial attention task). In order to manipulate exogenous spatial attention, each scene was preceded by a salient but task-irrelevant auditory stimulus (i.e., an auditory pre-cue) presented on the same side as the subsequent target ("cued" condition), on the opposite side to the subsequent target ("uncued" condition), or from both sides ("neutral" condition). Importantly, all these auditory pre-cues were "distracting" stimuli as they were entirely uninformative with regards to the location of the to-be-searched visual target, and thus completely useless to the accomplishment of the main task. To investigate the time course of the interaction between exogenous and endogenous attention, the SOA at which each auditory pre-cue was presented before the visual scene was varied across trials at one of three SOAs (either 50, 200, or 500 ms).

Using the just-mentioned design, we aimed to assess whether the endogenous control of spatial attention (as indexed by the performance at the visual search task) affected the resistance to crossmodal spatial distraction (i.e., the task-irrelevant exogenous pre-cues) differently across the different age groups and at the various SOAs. Based on the available literature showing a progressive increase in endogenous attention control from childhood to early adulthood (e.g., Enns et al., 1998; Rueda et al., 2004; Waszak et al., 2010; see also Cragg & Nation, 2008), and then a progressive decline in distraction inhibition from early adulthood through to old age (de Fockert et al., 2009; Healey et al., 2008), greater crossmodal spatial exogenous effects were expected in the two extreme age groups tested (i.e., young children, and older adults). Specifically, greater orienting (i.e., cued minus neutral trials) and reorienting (i.e., uncued minus neutral trials) effects were expected in young children and older adults as compared to young adults. Concerning the time-course of spatial exogenous distraction under endogenous control, on the basis of the previous literature (Castel, Chasteen, Scialfa, & Pratt, 2003; MacPherson et al., 2003), we would expect larger exogenous spatial distraction from task-irrelevant auditory pre-cues at shorter (50 ms) rather than at longer (200 and

500 ms) SOAs, and that this effect would be larger for the two extreme age groups tested here (young children and older adults), indexing a further reduction of endogenous attention when exogenous spatial distractors are delivered under increasing time pressure.

2. METHODS AND MATERIALS

2.1. Participants

A total of 124 participants volunteered for and took part in the study. Nine of these (four young children and five old children) did not complete the task, leaving a final sample of 115 participants, including: 23 young children attending the first grade and 32 older children attending the fifth grade of the "Istituto Comprensivo Perugia 7", San Sisto, Perugia, Italy, 30 young adults (university students) and 30 elderly adults recruited across several recreational centers for elderlies ("G. Balducci", "Maria Immacolata Gerini", & "Le Fonti") in Umbertide, Perugia, Italy. The participants' demographic characteristics for each group are summarized in Table 1. The appropriate sample size for this study was estimated with G*Power 3.1.9.2 (ANOVA, repeated measures, within-between interaction), taking into account: a medium-to-small effect size of 0.125 (predicted on the basis of a previous study using a similar design; Cavallina et al., 2018), a power of 85%, a significance level of 0.05, 4 groups, 9 measurements (i.e., 3 cue types x 3 SOAs), correlation among repeated measures of 0.5, and nonsphericity correction of 1. This indicated a minimum sample size of 23 participants per group. The exclusion criteria for children included diagnosis of neurodevelopmental disorders, as reported by parents or teachers. The exclusion criteria for the adults included the presence (self-reported) of any neurological, psychiatric, or cognitive disorder. All of the participants had normal or corrected-to-normal vision and normal hearing (the latter was tested prior to the start of the experiment, see "Stimuli and Task" section, below). The participants were naïve as to the main purpose of the study, which was conducted in accordance with the research ethics principles of the Declaration of Helsinki. All of the adult participants provided informed consent. Parental consent was obtained for each child who took part.

2.2. Stimuli and task

Each participant sat in a quiet room in front of a laptop computer. The laptop display was placed approximately 50 cm from the viewer (picture size = 29 x 22° of visual angle). A loudspeaker was placed on either side of the computer display. Before starting the experiment, participants' sound localization abilities were assessed by presenting them with a series of auditory stimuli (50 ms burst of white noise as in the main experiment; see below). This phase started with a subjective adjustment of the sound volume of the auditory stimulus as "clearly audible" (average = 62 dB; range 58-70 dB). Then, a series of twelve sounds was presented. The series included four stimuli presented from the left loudspeaker, four from the right, and four from both loudspeakers simultaneously (i.e., from a central location) (see Santangelo et al., 2007, for a similar experimental procedure). The sounds were presented in a random order and the participants had to indicate after each stimulus (with no time pressure) the location by pressing either the left-, right- or down-arrow (for central sounds) of the laptop keyboard. To participate in the main experiment, the participants had to achieve a minimum score of 80% (i.e., 10 out of 12 good responses). All participants exceeded this threshold, indicating a normal capability of sound localization of the current sample.

The visual stimuli used for the main – visual search – task consisted of cartoon scenes created using GoAnimate (GoAnimate © 2016; https://goanimate.com/) at a resolution of 1024 x 768 pixels (see Fig. 1A; cf. Cavallina et al., 2018). These scenes involved either internal (e.g., a kitchen, a bathroom, etc.) or external (e.g., a garden, a street, etc.) backgrounds. Each scene included several objects, but only one "agentive" element, that is, an animal (a dog, a cat, a rabbit, etc.) or a human (a young man or woman, etc.). These agentive elements were the targets to be searched for in the scenes. Targets were presented either at the left or right periphery of the visual scenes, at a mean eccentricity of 10.2 ± 2.3° of visual angle from the center of the display.

The visual search task comprised a sequence of 16 blocks of trials. At the beginning of each block, the target that was to be searched for in the following visual scenes was presented (see Fig. 1A). The target display was presented until the participant was convinced that he/she has memorized

the target identity and wished to continue. When ready, the participant pressed one of the two response keys ("D" or "L") to start the search task. After a blank display of 2000 ms, the first of nine visual scenes was presented. Each scene was displayed for 3000 ms, followed by an inter-stimulus interval (ITI) of 2000 ms, in which a blank screen was presented. The participants were required to discriminate the location of the target (i.e., the agentive element) within each scene and to press, as rapidly and accurately as possible, one of two response keys according to the target location, i.e., "D" for targets located on the left side of the scene (with respect to the central vertical meridian), and "L" for targets located on the right side of the scene. After all nine visual scenes had been displayed, a new block of trials began, starting with a new target display.

Importantly, each scene was preceded by the presentation of a burst of white noise on either the left, right, of both sides of the screen. The sound (duration 50 ms) was presented at one of three SOAs, i.e., either 50 ms (SOA50 condition), 200 ms (SOA200 condition), or 500 ms (SOA500 condition) before the presentation of the visual scene (Fig. 1B). The sound was equiprobably presented on either the left, right, or both sides of the scene, and the participants were informed that this sound was not informative about the target location. In cued trials, the sound was presented on the same side as the target; in uncued trials it was presented on the opposite side as the target in the scene; in neutral trials the sound was presented simultaneously on both sides of the scene, thus perceived as a central sound.

Each of the 16 blocks of trials included nine trials comprising all of the nine possible combinations between trial type and SOA (one cued, one neutral and one uncued trial at SOA50; one cued, one neutral and one uncued trial at SOA200; and one cued, one neutral and one uncued trial at SOA500). This resulted in a total of 144 trials. The order of blocks and the order of scenes and conditions within each block of trials was randomized across participants. Each scene was randomly assigned to cued, neutral, or uncued conditions, with the constraint that a cued, neutral, and uncued scene at the three different SOAs had to be presented within each block. Before the experimental session, participants practiced with a short training session comprising 4 blocks of 6 trials each, involving visual scenes not used in the main experiment.
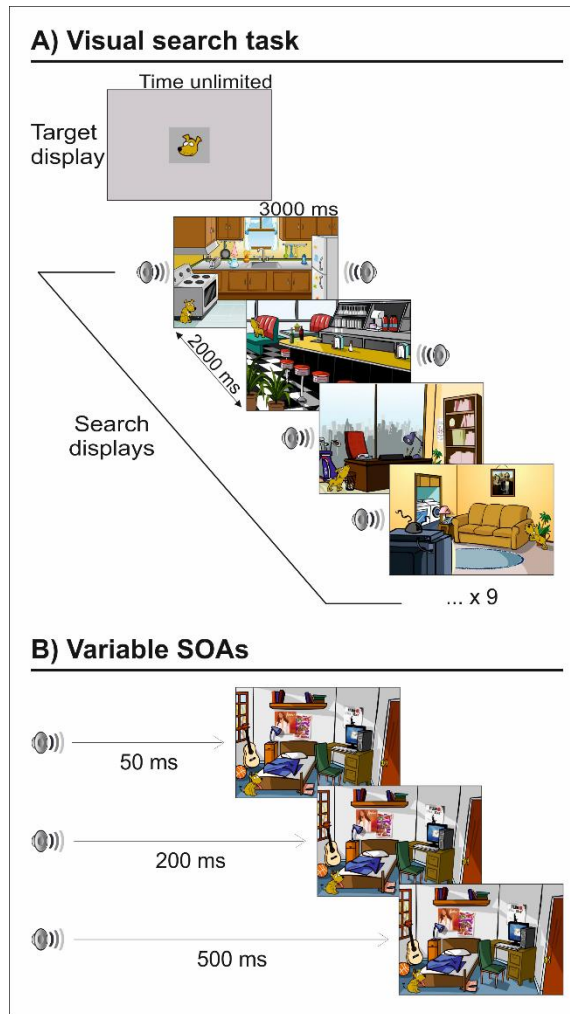
Fig 1. A) The sequence of events during a block of trials. This began with the presentation of a target display defining the to-be-searched-for-target in the following scenes. After any response button was pressed, a sequence of nine scenes was then presented. Each scene was displayed for 3000 ms, with an inter-stimulus interval of 2000 ms. Each picture was preceded by a 50 ms burst of white noise presented on either the left, right, or both hemifields, with a variable stimulus onset asynchrony (SOA), as illustrated in B. The participants pressed one of two response buttons according to the left vs. right location of the to-be-searched-for-target.

2.3. Data analysis

To assess whether search performance varied as a function of age group, trial type, and SOA, three generalized linear mixed models (GLMM) with maximum likelihood estimation using an adaptive Gauss–Hermite quadrature were used, one for each dependent variable. The dependent variables measured for each trial were: (i) the reaction time (RT), (ii) the accuracy value (i.e., the dichotomous correct/incorrect participant's response for that trial), and (iii) the normalized RTs (i.e., RT divided by the average RT for each participant, see below). The GLMM was selected because it

allowed us to model the participant's responses on a trial-by-trial basis, thus avoiding the use of a single summary measure for each participant (e.g., the mean or the median RT) while controlling for random effects of the participant. Moreover, the GLMM represents an extension of the linear mixed model (LMM). Whereas, in the LMM, the dependent variable needs to be transformed to meet the assumptions of normality, in the GLMM a non-normal distribution most appropriate to the distribution of the dependent variable could be directly specified into the model (e.g., Barr, Levy, Scheepers, & Tily, 2013; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). The RT data were modeled through an inverse Gaussian distribution with an inverse-square link function, which adequately fits the right-skewed distribution of the RT data (both at the level of the entire RT distribution and the single experimental conditions; see Baayen & Milin, 2010; Lo & Andrews, 2015). Accuracy data (correct or incorrect) were modeled through a binomial distribution with a logit link function (Dixon, 2008). Moreover, to investigate whether the observed effects on the RT were biased by a general increase in the response latencies in children and old adults compared to young adults, an additional analysis was run on normalized RTs. For this, the single-trial RTs were divided by the average RT of each participant (Matusz et al. 2015; see also, Huang-Pollock, Carr, & Nigg, 2002; Maylor & Lavie, 1998). As for the analysis of the RT, this analysis was also conducted with a GLMM assuming an inverse Gaussian distribution.

The only between-participants fixed effect of the three models was Age group (4 levels: young children, older children, young adults, or older adults). The within-participants fixed effects of the three models were Trial-type (3 levels: cued, neutral, uncued) and SOA (3 levels: 50, 200, 500 ms). The models also included all the relative interaction terms of the within- and between-participants fixed effects. Participants were added as a random intercept. Thus, the models for RT, normalized RT, and accuracy data were [RT ~ (Age group * Trial-type * SOA) + (1 | participants)]; [normalized RT ~ (Age group * Trial-type * SOA) + (1 | participants)] and [Accuracy ~ (Age group * Trial-type * SOA) + (1 | participants)], respectively. The GLMMs were performed using R (R Core Team, 2019, https://www.r-project.org/) and the models were run using the *glmer* command from the *lme4* package

(Bates, Mächler, Bolker, & Walker, 2015). For significance testing of fixed effect, the *Anova* function (using type III Wald chi-square test) from the *car* package (Fox, Friendly, & Weisberg, 2013) was used. Significant effects were compared using the *emmeans* package (Lenth, Singmann, Love, Buerkner, & Herve, 2018). This package determines whether there were significant differences between conditions based on the estimates and standard errors within the model. Reported p-values were corrected for multiple comparisons using the Tukey correction.

3. RESULTS

3.1. Reaction time

The GLMM on the RT revealed a significant main effect of Age group [X^2 (3, N = 115) = 1070.80, p < .001]. Post-hoc comparisons indicated faster RTs in the visual search task for the young adults (432 ms) than for older children (607 ms) and older adults (604 ms) (p < .001 for both), who did not significantly differ among themselves (compare the 2nd and the 4th bars in Fig. 2). These latter groups of participants, in turn, showed faster RTs than the young children (928 ms; p < .001 & p < .001, respectively) (see Fig. 2). The GLMM on the RTs also revealed a significant main effect of Trial-type [X^2 (2, N = 115) = 126.61, p < .001]. Post-hoc comparisons revealed that RTs were faster for cued (601 ms) than for neutral trials (623 ms; p < .001); in turn, performance for neutral trials was faster than for uncued trials (641 ms; p < .001). The main effect of SOA was also significant [X^2 (2, N = 115) = 248.65, p < .001], indicating faster search performance in the SOA500 condition (594 ms) than in the SOA200 condition (620 ms; p < .001), which, in turn, revealed faster performance than at the SOA50 condition (650 ms; p < .001).
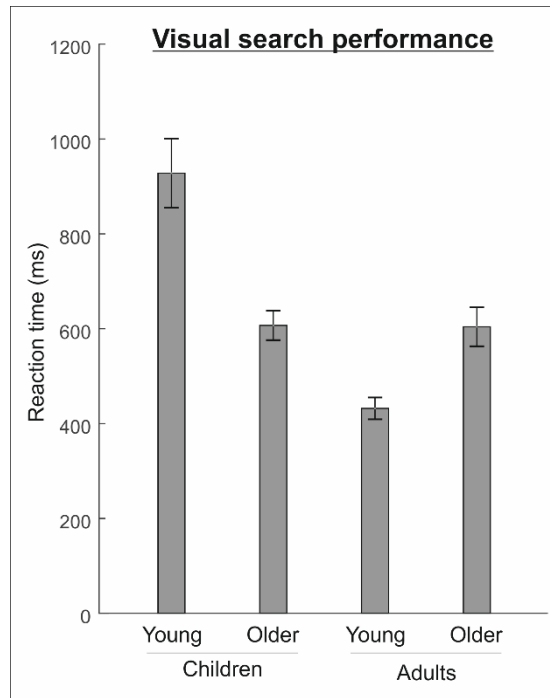
Fig 2. Mean reaction times (RT) in the visual search task across the four age groups. The error bars represent the standard error of the mean.

Although the three-way interaction between Age group, Trial-type, and SOA was not significant [X^2 (12, $N = 115$) = 15.60, $p = .21$], the model revealed both an Age group by Trial-type interaction [X^2 (6, $N = 115$) = 108.39, $p < .001$] and an Age group by SOA interaction [X^2 (6, N = 115) = 45.38, p < .001], indicating that the RTs of the different age groups were differently affected by both the cueing and the SOA manipulation. To avoid these above-mentioned effects being contaminated by the global decrease of response latencies of children and older adults (Matusz et al. 2015; see also, Huang-Pollock et al., 2002; Maylor & Lavie, 1998), a similar GLMM was conducted on normalized RTs (see above Data Analysis section). In this analysis, the RT of every trial was divided by the average RTs of each participant, thus removing the interference due to the subjective processing speed when trying to detect the orienting (i.e., cued minus neutral trials) and reorienting (uncued minus neutral trials) effect at any SOAs in the four age groups. This GLMM on the normalized RTs confirmed all of the results obtained in the previous GLMM on the RT, with the exception of the Age group effect, which was clearly not significant due to the normalization procedure (Trial-type: [X^2 (2, $N = 115$) = 185.56, $p < .001$]; SOA: [X^2 (2, $N = 115$) = 343.47, $p <$

.001]; Age group by Trial-type interaction: [X^2 (6, $N = 115$) = 92.48, $p < .001$]; Age group by SOA interaction: [X^2 (6, $N = 115$) = 48.07, $p < .001$]) (see Fig. 3 and Table 2). Subsequently, we conducted post-hoc comparisons aimed at investigating which age group showed significant spatial exogenous effects (i.e., faster RTs for cued than for uncued trials) and whether these effects were driven by the orienting (i.e., faster RTs for cued than for neutral trials) or reorienting (i.e., slower RTs for uncued than for neutral trials) sub-components of exogenous spatial attention. Moreover, we investigated whether the spatial exogenous effects were detectable only at specific SOAs (i.e., at the shortest SOA; Castel et al., 2003; MacPherson et al., 2003). The post-hoc analyses confirmed that the Trial-type differentially affected the normalized RT score depending on the age of the participants (see Fig. 3). Despite the RTs for uncued vs. cued trials being slowed in all age groups in terms of absolute values, these spatial exogenous effects were significant only for the young children and for the older adults. In young children, spatial exogenous effects were significant at the SOA50 (p = .017), without any further distinction in terms of orienting or reorienting sub-components of attention (see Fig. 3, top-left panel). In the older adults, spatial exogenous effects were significant at all SOAs (all p < .001, see Fig. 3, bottom-right panel). At SOA500, the effect was driven by a significant orienting component (p < .001); while at SOA200 and SOA50, the effect was driven by the orienting and reorienting sub-components of attention (orienting SOA200, p < .001, reorienting SOA200, p < .001, orienting SOA50, p = .018; reorienting SOA50, p < .001).
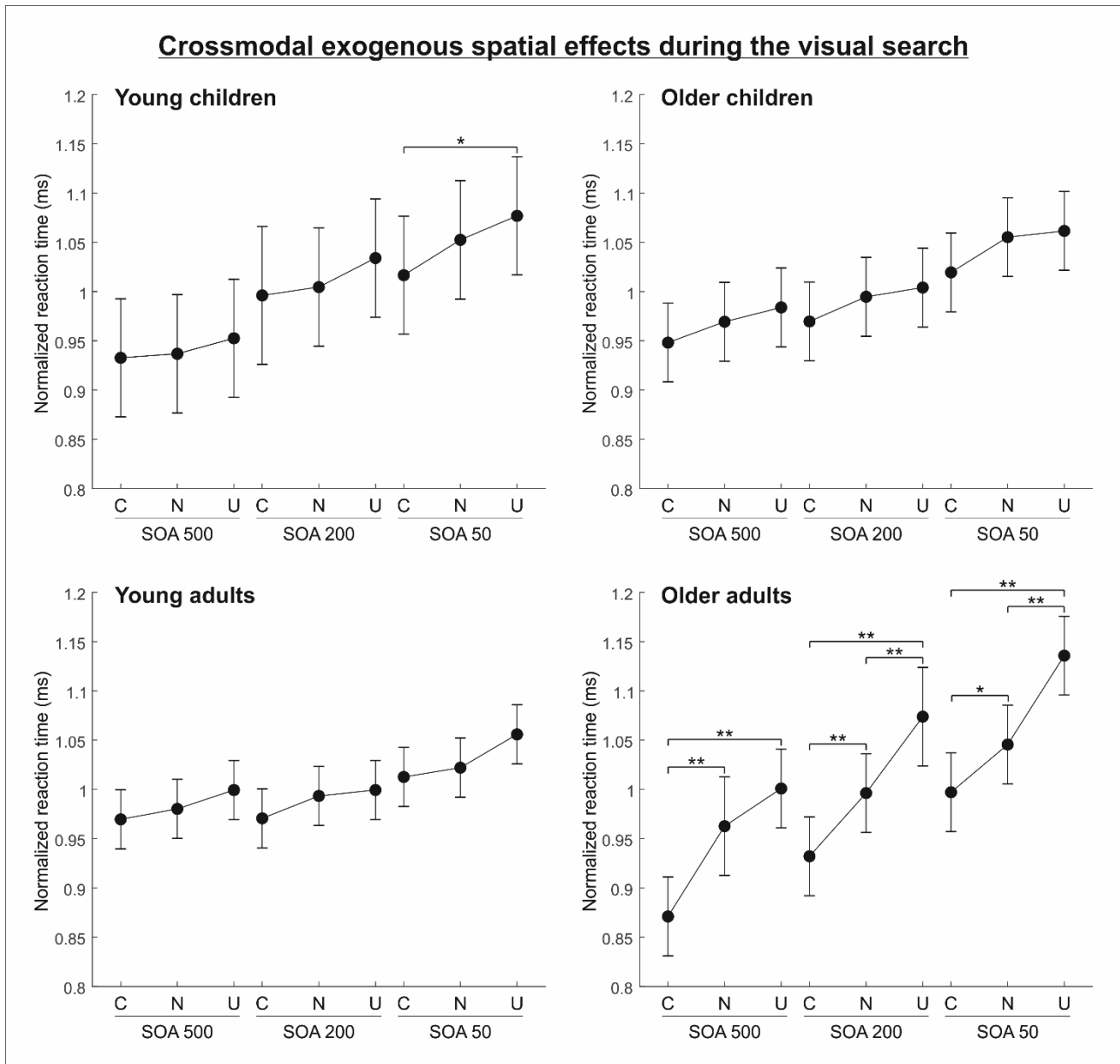
Fig 3. Mean normalized RT in the visual search task according to the trial type (cued, C; neutral, N; uncued, U) and SOAs (500, 200, 50 ms) across the four age groups. The error bars represent the standard error of the mean. * p < .05; ** p < .001

3.2. Accuracy

The GLMM on the accuracy data partially replicated the effect obtained with the RT (see Table 2). The analysis revealed a significant main effect of Age group [X^2 (3, N = 115) = 42.68, p < .001]. Post-hoc comparisons indicated lower accuracy at the visual search task in the young children compared with all the other groups (all ps < .001), which did not significantly differ from each other. The GLMM also revealed a significant main effect of Trial-type [X^2 (2, N = 115) = 58.99, p < .001].

Post-hoc comparisons revealed that the participants' responses were less accurate for uncued (90.16 ± 2.78%) than for neutral trials (92.7 ± 2.43%; p < .001), which, in turn, was less accurate than for cued trials (94.11 ± 2.2%; p < .007). The main effect of SOA was not significant [X^2 (2, $N = 115$) = 4.97, p = .083]. The model also revealed an Age group by Trial-type interaction [X^2 (6, $N = 115$) = 39.67, $p < .001$] and an Age group by SOA interaction [X^2 (6, N = 115) = 18.6, p = .005]. As in the RT analysis, we conducted post-hoc comparisons to investigate which age group showed accuracy-related spatial exogenous effects (i.e., worse performance with uncued than cued trials) and whether these spatial exogenous effects were significant only at certain SOAs. The post-hoc comparisons revealed that only older adults exhibited a significant reduction of accuracy for uncued vs. cued trials, and this was true for each SOA (all ps < .001).

DISCUSSION

The main goal of the present study was to investigate, in a single experimental task, the timecourse of the interplay between exogenous (i.e., the involuntary capture of attention by task-irrelevant auditory pre-cues) and endogenous spatial attention (i.e., a visual search task) from early childhood through to old age, and thereby to derive a complete picture of the lifespan development of endogenous resistance to crossmodal spatial exogenous distraction. To this end, young children (5-7 y.o.), older children (10-11 y.o.), young adults (20-27 y.o.), and older adults (62-86 y.o.) were presented with cartoon-like complex scenes in which their task was to search for a visual target. Each scene was preceded by a task-irrelevant sound presented either on the same or opposite side as the visual target, or simultaneously on both sides (cued, uncued, or neutral trials), at one of three SOAs (50, 200, or 500 ms).

Performance on the visual search task varied as a function of the participants' ages, with faster response time for young adults than for the older children and older adults, who, in turn, showed faster responses than the young children. In line with these results, young children also showed the worst performance in terms of accuracy as compared to all the other age groups. These findings

reinforce other findings published in the literature indicating that the control of endogenous attention follows a U-shaped developmental trajectory in which endogenous attention control increases from childhood to adulthood, followed by a decline into old age (de Fockert et al., 2009; Eckert, Keren, Roberts, Calhoun, & Harris, 2010; Enns et al., 1998; Hale, 1990; Healey et al., 2008; Rueda et al., 2004; Waszak et al., 2010). Our results also confirmed the capability of our exogenous manipulation since performance in the visual search task was better when the pre-cue was spatially aligned with the target (cued condition) than when it was neutral, which, in turn, elicited better performance than when the pre-cue was spatially misaligned with the target (uncued condition), irrespective of the age of the participants. These findings extend the notion that peripheral cues can "grab" - under the appropriate circumstances - spatial attention resources even under concurrent engagement of voluntary attention (Hillyard et al, 2016; Retsa, et al., 2020; Schreij, Owens, & Theeuwes, 2008; see also Santangelo & Spence, 2008), further clarifying the time course of this effect. Our results revealed a decrease of search performance as a function of decreasing SOA: the shorter the SOA, the harder participants found it to filter out the task-irrelevant auditory exogenous stimulus, irrespectively of the trial type and the participants' age. This finding highlights a general interference in the execution of the visual search task driven by task-irrelevant auditory stimuli (i.e., crossmodal exogenous distraction; Driver & Spence, 1998; Hughes, 2014; McDonald et al., 2003; McDonald & Ward, 2000; Spence & Driver, 2004; Van der Stoep et al., 2015), which becomes harder to control as the synchrony with the target presentation increases. It is well-known that the manipulation of SOA often results in a bi-phasic effect. In fact, SOAs shorter than 300 ms typically yield an increased benefit for cued (vs. uncued) trials (McDonald & Ward, 1999, 2000; Spence & McDonald, 2004; Van der Stoep et al., 2015). Conversely, SOAs longer than 300 ms yield poorer performance following cued (vs. uncued) trials (i.e., the IOR effect; see Klein, 2000; Klein & MacInnes, 1999). Although IOR was primarily observed in simple detection or discrimination tasks (see Klein, 2000; Lupiáñez, Martín-Arévalo, & Chica, 2013), IOR has also been reported during search in visual scenes (Klein & MacInnes, 1999; Smith & Henderson, 2011; see, for a review, Wang & Klein, 2010). However, with respect to this

previous literature already using visual scenes, here we manipulated spatial attentional orienting by means of auditory signals. With this experimental setting, we did not observe any IOR effect following the 500 ms SOA condition. Instead, we found that a SOA greater than 300 ms (i.e., the SOA500 condition) progressively reduced the impact of exogenous cues on the voluntary deployment of attention resource in the search task.

In line with the main goal of the present study, crossmodal exogenous attention capture under the concurrent engagement of endogenous spatial attention was shown to vary as a function of trial type ("cued", "neutral", and "uncued") and SOA in a different manner according to the participants' age. As predicted, consistent spatial exogenous effects in terms of RTs were found for the extreme ranges of the ages tested, i.e., in the young children and older adults, during the visual search task, indicating that, for these age groups, a difficulty coping with exogenous auditory distraction. In these two age groups, the interplay between exogenous distractors and the endogenous control of attention was affected by the SOA, but in a different manner. In the young children, significant spatial exogenous effects on the participants' RTs were found only at the shortest SOA (50 ms). This finding supports a susceptibility at this age to exogenous auditory stimuli, suggesting a reduced endogenous attention control in the presence of crossmodal exogenous distraction. These results are in line with other recent findings that revealed a greater impact of auditory distractors when attending to visual stimuli in young children as compared to adults (Matusz et al., 2015; Parker & Robinson, 2018; Robinson et al., 2018). This effect could be partially accounted for by a general "auditory dominance" in young children (as compared to the "visual dominance" normally seen at later stages of development), which might make young children more susceptible to auditory distraction compared with adults (Hirst, Cragg, & Allen, 2018; Hirst, Stacey, Cragg, Stacey, & Allen, 2018; Nava & Pavani, 2013; Robinson & Sloutsky, 2004, 2010; Sloutsky & Napolitano, 2003). The crossmodal distraction effect observed in the young children could also be interpreted in light of various accounts of multisensory development (Bahrick & Lickliter, 2012; Gibson, 1969; Lewkowicz, 2012; Maurer, Gibson, & Spector, 2012; Parker & Robinson, 2018), which propose that sensory modalities are more

strongly connected during early development (see Bremner, Lewkowicz & Spence, 2012). Distraction in young children is specifically limited to the short latency (50 ms SOA) condition. This might be interpreted as suggesting a particular role for greater connectivity between early stages of processing in sensory cortices (Maurer et al., 2012; although note that there is some skepticism concerning such accounts, Deroy & Spence, 2013). By contrast, the absence of distraction effects at longer SOAs in young children might be interpreted, as in the IOR effect, as a physiological disengagement from the spatial location signaled by an exogenous cue after an interval greater than 50 ms. Interestingly, at the 200 ms and 500 ms SOAs, no evidence was found of crossmodal spatial exogenous effects also for the other groups (except for the older adults; see the Discussion in the next paragraph), suggesting an overall efficiency in filtering out the task-irrelevant auditory cue and focus on the visual search task at these longer SOAs (i.e., resistance to distraction).

When considering the older adults, the picture changes entirely. For this group, we observed significant orienting and reorienting effects in terms of reaction time at each SOA (except for the reorienting component at the 500 ms SOA). Moreover, the accuracy data was also significantly affected by exogenous spatial distraction in this group at every SOA. In sum, this pattern of results highlights long-lasting benefits and persistent difficulties in filtering out task-irrelevant stimuli to focus on the searching task at each SOA. Overall, this indicates an extreme vulnerability to crossmodal exogenous distraction in aging (de Fockert et al., 2009; Healey et al., 2008; Langley et al., 2011). The existence of significant exogenous attentional effects in the older adults at each SOA, in terms of both reaction time and accuracy, is in line with a growing body of research revealing age-related impairments in other multisensory tasks (e.g., the "Temporal order judgment" task; Setti et al., 2014; Szymaszek, Sereda, Pöppel, & Szelag, 2009). In this task, two stimuli were presented in different sensory modalities (i.e., auditory and visual) with different ranges of SOAs. The participants were required to decide which of the two stimuli was presented first. With SOAs shorter than 70 ms, both young and old adults were not able to discriminate the temporal order in which the two stimuli were presented. This finding was interpreted by Setti and colleagues as a

tendency to merge the two sensory modalities into a unitary percept (see also Lewkowicz, 1996; Lewkowicz & Flom, 2014, for similar findings in early infancy). On the contrary, with longer SOAs (i.e., greater than 100 ms) old adults show a performance decline as compared to young adults, possibly because of the protracted processing of the first stimulus, which could extend the temporal window in which two stimuli are perceived as a unitary percept (Setti et al., 2011). Likewise, the exogenous effects observed here in aging at the longer SOA could originate from a protracted processing of the auditory distractor. This might have resulted in enhanced search performance when the auditory cue was spatially aligned with the visual target, but also in an increased interference with the voluntary disengagement of attentional resources when spatial reorienting was required for a target located on the opposite side of the scene (Castel et al., 2003).

The older children showed a pattern of performance that confirmed how this age group represented an intermediate stage of cognitive development. Indeed, their global processing speed was faster than that of the young children, but slower than young adults. Interestingly, in line with the U-shaped developmental pattern (de Fockert et al., 2009; Eckert et al., 2010; Enns et al., 1998; Hale, 1990; Healey et al., 2008; Rueda et al., 2004; Waszak et al., 2010), their performance was comparable with older adults (cf. the comparison between the 2nd and the 4th bars in Fig. 2). However, they showed better inhibition of crossmodal exogenous distraction as compared with the older adults, since no exogenous spatial effects were detected at any of the different SOAs.

Finally, the young adults did not show attentional effects at any SOA. These findings revealed no detectable interference from the auditory exogenous distractor in this age group overall, followed by a fast reallocation of endogenous attentional resources towards the to-be-searched visual target. Young adults therefore appeared to be most efficient in coping with crossmodal exogenous distraction when involved in a concurrent and demanding task requiring voluntary attention.

Overall, the current experiment revealed the existence of distinct timecourses of resistance to crossmodal exogenous distraction in young children and elderly adults. Several limitations should, however, also be taken into account in order to foster further research. First, we used "spatially"

distracting auditory cues without measuring their profile of phasic alerting. The inclusion of a "no auditory distraction" condition might allow future research to clarify the impact of the current auditory cues in increasing phasic alertness and readiness to respond to the upcoming event (i.e., the visual scene) in relation to the spatial and SOA manipulation (see, e.g., Petersen, Petersen, Bundesen, Vangkilde, & Habekost, 2017; Posner & Petersen, 1990). Second, recent studies suggest a number of factors, other than SOA, that could mitigate crossmodal exogenous distraction. For example, some studies have revealed that children appear to be paradoxically protected from interference by multisensory (e.g., audiovisual) distractors under certain conditions (e.g., when engaged in a high load task, Matusz et al., 2015, in accordance with the perceptual load theory, Lavie & Tsal, 1994; Murphy, Spence, & Dalton, 2017; or when audiovisual distractors are less familiar at a certain age, such as number words in contrast with digits, Matusz, Merkley, Faure, & Scerif, 2019). Moreover, in the presence of highly demanding endogenous attentional tasks, no differences between the effects of unisensory and multisensory exogenous cues were detected across three age groups of children with a different age range (Turoman et al., 2020). On the basis of this literature, it would be of interest to explore whether the crossmodal exogenous distraction observed in the current experiment would be reduced under greater demands of endogenous attention, e.g., by manipulating the perceptual load of the scenes. Third, we did not conduct a neuropsychological assessment of the elderly participants, to rule out, beyond the participants' own self-report, the lack of any cognitive impairments. Finally, we did not collect information related to level of education of our elderly participants, which might be an important factor for future research in this field to consider.

To conclude, our findings demonstrate different levels of efficiency in coping with crossmodal exogenous distraction under concurrent engagement of voluntary attention across age, which allowed us to draw a detailed picture of the developmental trajectory of the interplay between endogenous control and exogenous capture of attentional resources from childhood to aging.

Acknowledgments

Declarations of interest: None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at

https://data.mendeley.com/datasets/6vvpdtpvz4/draft?a=f2f8da1a-db2b-4d80-b131-21ec39a8f868.

References

Alain, C., & Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: Evidence for deficits in inhibitory control and sensory memory. *Psychology and Aging*, *14*, 507-519.

Andrés, P., Parmentier, F. B., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia*, *44*, 2564-2568.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3,* 12-28.

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183-206). Oxford University Press: Oxford, UK.

Barnhart, W. R., Rivera, S., & Robinson, C. W. (2018). Different patterns of modality dominance across development. *Acta Psychologica, 182,* 154-165.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255-278.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed models using lme4. *Journal of Statistical Software, 67,* 1-48.

Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General*, *134*, 207-221.

Bremner, A. J., Lewkowicz, D. J. & Spence, C. (Eds.). (2012). *Multisensory development*. Oxford University Press: Oxford, UK.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia*, *49*, 1446-1457.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*, 1860-1862.

Castel, A. D., Chasteen, A. L., Scialfa, C. T., & Pratt, J. (2003). Adult age differences in the time course of inhibition of return. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*, P256-P259.

Cavallina, C., Puccio, G., Capurso, M., Bremner, A. J., & Santangelo, V. (2018). Cognitive development attenuates audiovisual distraction and promotes the selection of task-relevant perceptual saliency during visual search on complex scenes. *Cognition*, *180*, 91-98.

Cragg, L., & Nation, K. (2008). Go or no-go? Developmental improvements in the efficiency of response inhibition in mid-childhood. *Developmental Science, 11*, 819-827.

de Fockert, J. W., Ramchurn, A., Van Velzen, J., Bergström, Z., & Bunce, D. (2009). Behavioral and ERP evidence of greater distractor processing in old age. *Brain Research*, *1282*, 67-73.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45-75.

Deroy, O., & Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience & Biobehavioral Reviews*, *37*, 1240-1253.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59,* 447-456.

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*, 254-262.

Eckert, M. A., Keren, N. I., Roberts, D. R., Calhoun, V. D., & Harris, K. C. (2010). Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*, *4*:10.

Enns, J. T., Brodeur, D. A., & Trick, L. M. (1998). Selective attention over the life span: Behavioral measures. In J. E. Richards (Ed.), *Cognitive neuroscience of attention: A developmental perspective* (pp. 393-418). Mahwah, NJ: Lawrence Erlbaum Associates.

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience*, 8, 1643-1646.

Folk, C. L., & Hoyer, W. J. (1992). Aging and shifts of visual spatial attention. *Psychology and Aging*, *7*, 453-465.

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, *111*, 345-355.

Fox, J., Friendly, M., & Weisberg, S. (2013). Hypothesis tests for multivariate linear models using the car package. *The R Journal, 5,* 39-52.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts: New York, NY.

Guerreiro, M. J., Murphy, D. R., & Van Gerven, P. W. (2013). Making sense of age-related distractibility: The critical role of sensory modality. *Acta Psychologica*, *142*, 184-194.

Guerreiro, M. J., & Van Gerven, P. W. (2011). Now you see it, now you don't: Evidence for age-dependent and age-independent cross-modal distraction. *Psychology and Aging*, *26*, 415-426.

Gumenyuk, V., Korzyukov, O., Alho, K., Escera, C., & Näätänen, R. (2004). Effects of auditory distraction on electrophysiological brain activity and performance in children aged 8–13 years. *Psychophysiology*, *41*, 30-36.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development*, *61*, 653-663.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175-204). Cambridge, MA: Academic Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning & Motivation*, *22*, 193-225.

Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. *Progress in Brain Research*, *169*, 353-363.

Hillyard, S. A., Störmer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia, 83,* 170-178.

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience & Biobehavioral Reviews, 94,* 286-301.

Hirst, R. J., Stacey, J. E., Cragg, L., Stacey, P. C., & Allen, H. A. (2018). The threshold for the McGurk effect in audio-visual noise decreases with development. *Scientific Reports, 8:*1-12.

Huang-Pollock, C. L., Carr, T. H., & Nigg, J. T. (2002). Development of selective attention: Perceptual load influences early versus late attentional selection in children and adults. *Developmental Psychology, 38,* 363–375.

Hugenschmidt, C. E., Peiffer, A. M., McCoy, T. P., Hayasaka, S., & Laurienti, P. J. (2009). Preservation of crossmodal selective attention in healthy aging. *Experimental Brain Research, 198,* 273-285.

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*, 30-41.

Iarocci, G., Enns, J. T., Randolph, B., & Burack, J. A. (2009). The modulation of visual orienting reflexes across the lifespan. *Developmental Science*, *12*, 715-724.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4,* 138-147.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*, 346-352.

Langley, L. K., Friesen, C. K., Saville, A. L., & Ciernia, A. T. (2011). Timing of reflexive visuospatial orienting in young, young-old, and old-old adults. *Attention, Perception, & Psychophysics*, *73*, 1546-1561.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339-354.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56,* 183-197.

Leclercq, V., & Siéroff, E. (2013). Development of endogenous orienting of attention in school-age children. *Child Neuropsychology*, *19*, 400-419.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version, 1*:3.

Lewkowicz, D. J. (1996). Perception of auditory-visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1094-1106.

Lewkowicz, D. J. (2012). The unexpected effects of experience on the development of multisensory perception in primates. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 159-182). Oxford University Press: Oxford, UK.

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Development, 85*, 685-694.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*:1171.

Lupiáñez, J., Klein, R. M., & Bartolomeo, P. (2006). Inhibition of return: Twenty years after. *Cognitive Neuropsychology*, *23*, 1003-1014.

Lupiáñez, J., Martín-Arévalo, E., & Chica, A. B. (2013). Is inhibition of return due to attentional disengagement or to a detection cost? The detection cost theory of IOR. *Psicologica: International Journal of Methodology and Experimental Psychology*, *34*, 221-252.

Macaluso, E., & Doricchi, F. (2013). Attention and predictions: Control of spatial attention beyond the endogenous-exogenous dichotomy. *Frontiers in Human Neuroscience*, *7*:685.

MacPherson, A. C., Klein, R. M., & Moore, C. (2003). Inhibition of return in children and adolescents. *Journal of Experimental Child Psychology*, *85*, 337-351.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94,* 305-315.

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.

Matusz, P. J., Merkley, R., Faure, M., & Scerif, G. (2019). Expert attention: Attentional allocation depends on the differential development of multisensory number representations. *Cognition, 186,* 171-177.

Maurer, D., Gibson, L. C., & Spector, F. (2012). Infant synaesthesia. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 229-250). Oxford University Press: Oxford, UK.

Maylor, E. A., & Lavie, N. (1998). The influence of perceptual load on age differences in selective attention. *Psychology and Aging, 13,* 563-573.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, *15*, 10-19.

McDonald, J. J., & Ward, L. M. (1999). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1234-1252.

McDonald, J. J., & Ward, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, *11*, 167-171.

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research, 352,* 40-48.

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: Converging evidence from the Colavita effect and the sound-induced flash illusion. *Child Development, 84,* 604-616.

Parker, J. L., & Robinson, C. W. (2018). Changes in multisensory integration across the lifespan. *Psychology and Aging, 33,* 545-558.

Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences*, *18*, 242-250.

Petersen, A., Petersen, A. H., Bundesen, C., Vangkilde, S., & Habekost, T. (2017). The effect of phasic auditory alerting on visual perception. *Cognition, 165,* 73-81.

Pierce, A. M., McDonald, J. J., & Green, J. J. (2018). Electrophysiological evidence of an attentional bias in crossmodal inhibition of return. *Neuropsychologia*, *114*, 11-18.

Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica, 86,* 227-272.

Poliakoff, E., Ashworth, S., Lowe, C., & Spence, C. (2006). Vision and touch in ageing: Crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia, 44,* 507-517.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention & Performance X: Control of Language Processes*, *32*, 531-556.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25-42.

Retsa, C., Matusz, P. J., Schnupp, J. W., & Murray, M. M. (2020). Selective attention to sound features mediates cross-modal activation of visual cortices. *Neuropsychologia, 144*:107498.

Robinson, C. W., Hawthorn, A. M., & Rahman, A. N. (2018). Developmental differences in filtering auditory and visual distractors during visual selective attention. *Frontiers in Psychology, 9*:2564.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75,* 1387-1401.

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 135-141.

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*, 1029-1040.

Santangelo, V., & Macaluso, E. (2012). Spatial attention and audiovisual processing. In B. E. Stein (Ed.), *The New Handbook of Multisensory Processing* (pp. 359-370). Cambridge, MA: MIT Press.

Santangelo, V., Olivetti Belardinelli, M., & Spence, C. (2007). The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 137-148.

Santangelo, V., Olivetti Belardinelli, M., Spence, C., & Macaluso, E. (2009). Interactions between voluntary and stimulus-driven spatial attention mechanisms across sensory modalities. *Journal of Cognitive Neuroscience*, *21*, 2384-2397.

Santangelo, V., & Spence, C. (2008). Is the exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Consciousness & Cognition*, *17*, 989-1015.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception & Psychophysics*, *70*, 208-218.

Setti, A., Finnigan, S., Sobolewski, R., McLaren, L., Robertson, I. H., Reilly, R. B., Kenny, R. A., & Newell, F. N. (2011). Audiovisual temporal discrimination is less efficient with aging: An event-related potential study. *Neuroreport, 22,* 554-558.

Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R. A., & Newell, F. N. (2014). Improving the efficiency of multisensory integration in older adults: Audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. *Neuropsychologia, 61,* 259-268.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74,* 822-833.

Smith, T. J., & Henderson, J. M. (2011). Looking back at Waldo: Oculomotor inhibition of return does not prevent return fixations. *Journal of Vision, 11*, 1-11.

Spalek, T. M., Falcon, L. J., & Di Lollo, V. (2006). Attentional blink and attentional capture: Endogenous versus exogenous control over paying attention to two important events in close succession. *Perception & Psychophysics*, *68*, 674-684.

Spence, C., & Driver, J. (1998). Auditory and audiovisual inhibition of return. *Perception & Psychophysics*, *60*, 125-139.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford, UK: Oxford University Press.

Spence, C., Lloyd, D., McGlone, F., Nicholls, M. E. R., & Driver, J. (2000). Inhibition of return is supramodal: A demonstration between all possible pairings of vision, touch and audition. *Experimental Brain Research*, *134*, 42-48.

Spence, C., & McDonald, J. (2004). The crossmodal consequences of the exogenous spatial orienting of attention In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 3-25). Cambridge, MA: MIT Press.

Spence, C., & Soto-Faraco, S. (2020). Crossmodal attention applied: Lessons for/from driving. *Cambridge elements of attention*. Cambridge, UK: Cambridge University Press.

Störmer, V. S. (2019). Orienting spatial attention to sounds enhances visual processing. *Current Opinion in Psychology, 29,* 193-198.

Szymaszek, A., Sereda, M., Pöppel, E., & Szelag, E. (2009). Individual differences in the perception of temporal order: The effect of age and cognition. *Cognitive Neuropsychology, 26,* 135-147.

Tassinari, G., & Campara, D. (1996). Consequences of covert orienting to non-informative stimuli of different modalities: A unitary mechanism? *Neuropsychologia*, *34*, 235-245.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*, 77-99.

Tipper, S. P. (1991). Less attentional selectivity as a result of declining inhibition in older adults. *Bulletin of the Psychonomic Society*, *29*, 45-47.

Trick, L. M., & Enns, J. T. (1998). Lifespan changes in attention: The visual search task. *Cognitive Development*, *13*, 369-386.

Turoman, N., Tivadar, R. I., Retsa, C., Maillard, A. M., Scerif, G., & Matusz, P. J. (2020). The development of attentional control mechanisms in multisensory environments. *BioRxiv.*

Van der Lubbe, R. H., & Postma, A. (2005). Interruption from irrelevant auditory and visual onsets even when attention is in a focused state. *Experimental Brain Research, 164,* 464-471.

Van der Stoep, N., Nijboer, T. C. W., Van der Stigchel, S., & Spence, C. (2015). Multisensory interactions in the depth plane in front and rear space: A review. *Neuropsychologia*, *70*, 335-349.

Van Gerven, P. W., & Guerreiro, M. J. (2016). Selective attention and sensory modality in aging: Curses and blessings. *Frontiers in Human Neuroscience, 10*:147.

Van Gerven, P. W., & Murphy, D. R. (2010). Aging and distraction by irrelevant speech: Does emotional valence matter? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*, 667-670.

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research, 50*, 220-228.

Waszak, F., Li, S. C., & Hommel, B. (2010). The development of attentional networks: Cross-sectional findings from a life span sample. *Developmental Psychology*, *46*, 337-349.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 121-134.

Table 1. Participants' demographic characteristics for each age group.

	Number of participants	Gender		Mean age ± S. D.		Range	
		Males	Females	Years	Months	Years	Months
Young children	23	12	11	6.2 ± 0.5	74.4 ± 6.0	5-7	60-84
Older children	32	16	16	10.2 ± 0.4	122.4 ± 4.8	10-11	120-132
Young adults	30	7	23	22.2 ± 1.7	266.4 ± 20.4	20-27	240-324
Older adults	30	15	15	71.7 ± 5.3	860.4 ± 63.6	62-86	744-1032

Table 2. Reaction time (RT), normalized RT (nRT), and accuracy (ACC) ± standard error of the mean (in brackets) at the visual search task according to trial type (cued, neutral, and uncued) and SOAs (500, 200, and 50 ms) in the four age groups.

	Young children			Older children			Young adults			Older adults		
	Cued	Neutral	Uncued	Cued	Neutral	Uncued	Cued	Neutral	Uncued	Cued	Neutral	Uncued
RT (ms)												
SOA500	866(70)	872(74)	881(66)	576(29)	586(28)	597(31)	420(23)	423(21)	430(21)	525(37)	585(44)	610(43)
SOA200	918(72)	934(76)	955(73)	590(31)	606(32)	610(29)	419(22)	430(23)	432(23)	562(40)	602(39)	653(45)
SOA50	943(71)	975(75)	1005(76)	618(31)	642(32)	646(31)	438(24)	441(22)	457(23)	598(36)	632(39)	680(40)
nRT (ms)												
SOA500	0.93(0.06)	0.94(0.06)	0.95(0.06)	0.95(0.04)	0.97(0.04)	0.98(0.04)	0.97(0.03)	0.98(0.03)	1.00(0.03)	0.87(0.04)	0.96(0.05)	1.00(0.04)
SOA200	1.00(0.07)	1.00(0.06)	1.03(0.06)	0.97(0.04)	0.99(0.04)	1.00(0.04)	0.97(0.03)	0.99(0.03)	1.00(0.03)	0.93(0.04)	1.00(0.04)	1.07(0.05)
SOA50	1.02(0.06)	1.05(0.06)	1.08(0.06)	1.02(0.04)	1.06(0.04)	1.06(0.04)	1.01(0.03)	1.02(0.03)	1.06(0.03)	1.00(0.04)	1.05(0.04)	1.14(0.04)
ACC (%)												
SOA500	88.3(6.7)	89.1(6.5)	88.6(6.6)	95.1(3.8)	96.1(3.4)	92.6(4.6)	96.7(3.3)	95.2(3.9)	94.4(4.2)	95.0(4.0)	94.0(4.4)	86.9(6.2)
SOA200	89.7(6.4)	88.6(6.6)	87.2(7)	94.3(4.1)	91.0(5.1)	91.8(4.9)	95.2(3.9)	93.5(4.5)	95.2(3.9)	96.9(3.2)	93.8(4.4)	87.5(6.0)
SOA50	89.9(6.3)	88(6.8)	86.4(7.2)	92.8(4.6)	91.8(4.9)	88.9(5.6)	95.8(3.7)	92.3(4.9)	91.7(5.1)	96.3(3.5)	96.0(3.6)	88.8(5.8)