Stroop Effects in Alzheimer's Disease: Selective Attention Speed of Processing, or Color-naming? A Meta-Analysis

Boaz M. Ben-David[a,b,c,d,*], Anita Tewari[b], Vered Shakuf[a] and Pascal H.H.M. Van Lieshout[b,c,d,e,f]

[a]Communication, Aging and Neuropsychology Lab (CANlab), School of Psychology, Interdisciplinary Center (IDC), Herzliya, Israel
[b]Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, Canada
[c]Toronto Rehabilitation Institute, Canada
[d]Graduate Department of Rehabilitation Sciences, University of Toronto, Canada
[e]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Canada
[f]Department of Psychology, University of Toronto, Canada

Abstract. Selective attention, an essential part of daily activity, is often impaired in people with Alzheimer's disease (AD). Usually, it is measured by the color-word Stroop test. However, there is no universal agreement whether performance on the Stroop task changes significantly in AD patients; or if so, whether an increase in Stroop effects reflects a decrease in selective attention, a slowing in generalized speed of processing (SOP), or is the result of degraded color-vision. The current study investigated the impact of AD on Stroop performance and its potential sources in a meta-analysis and mathematical modeling of 18 studies, comparing 637 AD patients with 977 healthy age-matched participants. We found a significant increase in Stroop effects for AD patients, across studies. This AD-related change was associated with a slowing in SOP. However, after correcting for a bias in the distribution of latencies, SOP could only explain a moderate portion of the total variance (25%). Moreover, we found strong evidence for an AD-related increase in the latency difference between naming the font-color and reading color-neutral stimuli ($r^2 = 0.98$). This increase in the dimensional imbalance between color-naming and word-reading was found to explain a significant portion of the AD-related increase in Stroop effects ($r^2 = 0.87$), hinting on a possible sensory source. In conclusion, our analysis highlights the importance of controlling for sensory degradation and SOP when testing cognitive performance and, specifically, selective attention in AD patients. We also suggest possible measures and tools to better test for selective attention in AD.

Keywords: Aging, Alzheimer's disease, color vision, selective attention, sensory and cognitive interaction, speed of processing, Stroop test, visual perception

INTRODUCTION

Attending selectively to certain aspects in a scene, while ignoring or actively suppressing others, is an everyday occurrence. For example, when crossing the street in a busy intersection, one has to spot oncoming cars while ignoring distracting information such as the color of the cars or other pedestrians. In general, older adults are less efficient at exercising selective attention. This is often suggested as the possible source of age-related reduction in performance in complex environments (such as listening in adverse situations [1–3]). To date, the most widely used test for assessing selective attention in older individuals, both in clinical and experimental settings, is the color-word Stroop

*Correspondence to: Boaz M. Ben-David, PhD, Communication Aging and Neuropsychology lab (CAN lab), School of Psychology, Interdisciplinary Center (IDC) Herzliya, P.O. Box 167, Herzliya 4610101, Israel. Tel.: +972 9 9602429; Fax: +972 9 9602845; E-mail: boaz.ben.david@idc.ac.il.

task ([4]; for relevant reviews, see [5, 6]). The Stroop test is a part of neuropsychological screening batteries to assess the severity of performance deficiencies in pathological populations and to thereby target the correct clinical intervention [7–9]). Recently, studies have raised doubts on the external validity of the Stroop test, and its relevance to assess cognitive function in pathological [6, 10] and older populations [11]. The current paper links this literature to Alzheimer's disease (AD), trying to assess whether the commonly used Stroop test can serve as a true yardstick to measure selective attention in this population. In addition, we wish to explore if the Stroop test underestimates the cognitive abilities of AD patients, biasing the evaluation of these individuals. As a first step, we present a review of the test.

In the classic version of the color-word Stroop task, participants are asked to name the font color of stimuli; some are colored color-names for which the font color is incongruent with their content (e.g., RED printed in blue) and some carry a (color-) neutral content (e.g., a colored string of X's). Generally, for both younger and older individuals, a Stroop Interference (SI) ensues, that is, color-naming incongruent stimuli (Ci) takes longer than color-naming neutral stimuli (Cn),

$$SI = Ci - Cn. \tag{1}$$

The magnitude of SI has been consistently found to be larger for healthy older adults and traditionally taken to reflect a deficiency in selective attention, generated by an age-related deficit in inhibition processes [12, 13]. However, other sources have been suggested for this increase in SI, unrelated to selective attention. In particular, a general slowing in speed of processing [14] and, more recently, a sensory deficit [6, 10] have been discussed in the literature. Alternatively, it could reflect a combination of these three sources [11].

Alzheimer's disease and Stroop effects

Stroop performance has often been found to be further impaired for older individuals with deteriorated cognitive functions; specifically, for older adults diagnosed with symptoms on the Dementia spectrum [15], a cognitive decline of chronic or progressive nature that produces an appreciable decline in intellectual functioning, usually while interfering with personal activities of daily living [16]. Dementia is typically associated with the following conditions: vascular dementia, Lewy body dementia, frontotemporal dementia, and AD. Impaired attentional performance leads to emotional dissatisfaction and frustration, ulti-

mately placing dementia patients at a higher risk for depression [17]. Correctly assessing limitations in selective attention is therefore an important step in addressing the difficulties dementia patients encounter throughout life. To focus this study, we limit the scope of the analysis to the most common type of dementia—AD.

Past findings

AD is one of the most common causes of cognitive deficits that affect older adults. It is estimated that more than 26.6 million individuals are afflicted with AD worldwide [18]. Symptoms related to AD include memory impairment, a decline in learning ability, disorientation, and other behavioral and cognitive changes. The disease is progressive, and associated symptoms are known to worsen over time [19, 20]. The majority of the literature reports an AD-related increase in Ci and SI values (compared to healthy older adults). This is illustrated in the latest review on Stroop and AD by Amieva and colleagues [15] conducted a decade ago. In an analysis of six studies[1], Amieva et al. found that the AD-related increase in incongruent trials (Ci) was significantly larger than for neutral trails (Cn), resulting in a significant increase in Stroop effects (SI, see Eq. 1) for AD patients. The authors concluded that this specific deficit in Ci reflects an AD-related reduction in the ability to inhibit processing (the meaning of) the printed word. Amieva's meta-analysis stands in contrast to previous findings in the literature. Forty years ago, in one of the earliest studies examining Stroop in older adults with impaired cognitive functions (then referred to as 'Organic Brain Syndrome', a somewhat outdated term that includes AD, [21]), Bettner and colleagues [22] reported that the disease had no significant impact on Stroop performance (specifically, Ci). More recently, a literature review on Stroop and AD by Spieler et al. [23] found no AD-related increase in SI after controlling for general slowing associated with AD (in a mathematical modeling of the response-latency distribution). These discrepancies in the literature call for a fresh examination of the Stroop test in AD. This is especially relevant since the meta-analysis by Amieva et al. conducted a decade ago examined only seven Stroop and AD studies, and since its publication, many other studies have been published. In the current study, we provide a new examination of the data on Stroop and dementia,

---

[1]Note, one study was excluded from the Brinley analysis presented in the study by Amieva and colleagues [15].

combining studies examined by Amieva et al. and studies published since Amieva's meta-analysis. Furthermore, alongside the traditional analysis on Stroop effects, we examine more specifically three possible sources for the AD-related increase in Stroop effects presented in the following section.

Three possible hypotheses for AD-related increase in Stroop effects

The literature on healthy aging suggests three possible hypotheses for age-related changes in Stroop. These hypotheses have also been employed to explain differences in Stroop performance between healthy older adults and AD patients: 1) a robust reduction in selective attention in AD [15]; 2) an increase in generalized slowing of processing in AD above age-related effects [24]; and 3) a sensory deficit that carries a specific impact on color-naming in unhealthy aging [25]. The current study presents the first systematic comparison of the potential explanatory strength of these possibilities in the realm of AD, attempting to uncover the source of an AD-related deficit in Stroop performance. In the following sections, we present each of these hypotheses in more detail and explore the evidence that exists in the current literature to support them.

Can differences in Stroop interference (SI) reflect AD-related differences in selective attention?

Generally, performance on the Stroop task is taken as the "gold standard" for selective attention [26]. Recall that in Ci, the participant is asked to ignore the (incongruent) lexical content of the stimulus, and focus only on its font color. In Cn, the neutral lexical stimulus does not carry any semantic content that may interfere with color-naming. Thus, the latency difference between the two, SI (Stroop Interference), reflects the additional cost associated with ignoring the lexical content, and an increase in the magnitude of SI suggests a decrease in the ability to selectively inhibit lexical processing. Indeed, many of the studies that report a significant AD-related increase in SI attribute it to a reduction in selective attention following AD.

AD has been commonly related to impairments in executive functions, a general term for cognitive processes including planning, working memory, and selective attention [27]. Specifically, the literature points to selective attention (as measured by the Stroop task) as one of the major areas of cognitive impairments in AD [28]. Following this, the reduction in Stroop

performance reflects the patients' impaired ability to inhibit more automatic responses (such as reading the printed color-words in the Stroop context, [29]), or their reduced ability to disengage and shift attention from one aspect of a task to another (e.g., from the distracting lexical dimension to the color dimension, [30]). This attentional impairment was found even in earlier stages of AD [31, 32], and may be related to damage to frontal lobe regions (linked with selective attention; e.g., ACC [33, 34]), which starts in very early stages of AD [35]. To address this issue, the current study will first test for a significant AD-related increase in Stroop effects, across studies and laboratories. Next, we will test for other possible mediating factors, as discussed below.

Can differences in SI reflect AD-related changes in cognitive slowing?

Cognitive slowing, a characteristic of the normal aging process [36–38], is often considered a source of deteriorated cognitive performance with normal aging and AD patients alike. In a generalized cognitive slowing framework, reaction times (RTs) on all tasks increase to the same extent [39], that is,

$$RT_{AD} = a + b^* RT_{Control} \qquad (b > 1), \qquad (2)$$

Where *b* reflects generalized slowing of central cognitive processes. Indeed, several meta-analyses (for a recent review, see [40]) confirmed age-related changes predicted by this relationship, using a Brinley [41] analysis. In this type of analysis, latencies for older adults on a task are plotted as a function of latencies for younger adults on the same task, across several studies and tasks. Generalized speed of processing is supported if a single regression line with a single slope value, *b*, connects latencies for older and younger adults across all tasks (but see a critique on the Brinley analysis in [42]).

Nebes and Brady [24] (see also [43]), in their meta-analysis of 55 experiments, suggested that cognitive slowing for AD patients is greater than non-pathological age-related slowing. In accordance with these claims, Brinley's analysis, which was conducted separately for healthy older adults and for AD patients (as a function of latencies for younger adults), revealed a larger slope (*b* factor) for AD patients (2.3) than for healthy older adults (1.4). These linear functions were found to explain up to 83% of the variance in response times across the different tasks. Given these findings on cognitive slowing in AD patients, it is highly possible that such slowing underlies AD-related increases

in SI as well [44]. Thus, if Eq. 2 holds for both tasks with the same variables, then

$$
\begin{cases}
SI_{Control} = Ci_{Control} - Cn_{Control} \\
SI_{AD} = Ci_{AD} - Cn_{AD} = (a + b^*Ci_{control}) \\
\qquad - (a + b^*Cn_{control}) \Rightarrow \\
SI_{AD} = b^*(Ci_{Control} - Cn_{Control}) = b^*SI_{Control}
\end{cases}
\tag{3}
$$

That is, if AD leads to an increased slowing across the different tasks above and beyond an age-related impact ($b > 1$), a larger SI for AD patients ensues, irrespective of changes in selective attention. On the other hand, if there is a real AD-related difference in selective attention irrespective of speed of processing (SOP), the slope for an increase in incongruent trials, Ci, as a function of AD, should be significantly larger than the slope for neutral ones, Cn ($b_{Ci} > b_{Cn}$). Can the observed AD-related slowing in SOP underlie the differences in Stroop performance? The evidence so far does not support this notion. Amieva and colleagues [15], in their meta-analysis, found the AD-related increase in SI to persist even after controlling for SOP, concluding that: "Alzheimer's disease has a differential effect on interference [Ci] as opposed to baseline latencies [Cn], supporting the idea of a specific inhibitory deficit separate from general slowing" (p. 956, the terminology used in our study was added in brackets). Similarly, in a visual-search study [45], AD-related slowing was only observed in a search that involves selective attention (conjunctive search). No differences were observed in a simple search task that does not require selective attention, leading the authors to conclude that "selective attention is affected in AD patients, above and beyond the degree of impairment expected according to the notion of generalized cognitive slowing" (p. 241).

In sum, there are contradicting findings on the possible role of SOP as a mediator for AD-related changes in selective attention. The current meta-analysis re-examines SOP as a possible underlying factor for changes in Stroop performance for AD patients, using a relatively large number of studies (combining past and recent AD and Stroop studies). The next section introduces a third and novel line of explanation for an AD-related decline in Stroop performance, linking sensory degradation with cognitive performance.

Can differences in Stroop performance reflect AD-related changes in color-vision?

Recently, changes in selective attention performance across the lifespan have been related to sensory degradation. Specifically, color vision deteriorates rapidly after the age of 55 (related to the yellowing of the lens and neural degeneration [46, 47]), but on the other hand, reading is relatively preserved in aging (given good legibility [48, 49]). This disproportionate age-related slowing in color-naming, versus a moderate age-related change in reading, is gauged by the Dimensional Imbalance (DI) measure: The latency difference between naming the font color of a neutral stimulus (e.g., responding "red" to a red string of X's, Cn) and reading aloud a color word printed black-on-white (e.g., responding "red" to the word RED printed in black, Rn [48, 50]). In the Tectonic Theory of the Stroop phenomenon, Melara and Algom [51] proposed that this type of increase in DI might by itself inflate Stroop effects; the greater the imbalance, the greater the SI.

The link between DI and Stroop effects is supported by three separate meta-analyses which found a significant positive correlation between DI and Ci in different populations: healthy younger adults [51], healthy older adults [6], and traumatic brain injury patients [11]. Similarly, direct experimental manipulations of DI show that Stroop effects are sensitive to psychophysical characteristics of the stimuli. Decreasing the DI (by minimizing font size—inflating latencies for reading, Rn), generated a decrease [52] or elimination [53] of Stroop effects. Increasing the DI (by rendering colors difficult to distinguish– increasing the Cn) resulted in an increase of Stroop effects [10]. In other words, the imbalance between fast reading and slower color-naming can engender a Stroop effect. As this imbalance increases (e.g., by disproportionally slowing color-naming), so do Stroop effects.

Some evidence in the literature suggests that AD might inflate DI to a larger extent than healthy aging (uniquely impairing color-vision), and as a consequence, might contribute to the increase in Stroop effects. A series of studies by Cronin-Golomb and her colleagues [54–57] found that AD patients have diminished color-vision, in addition to other observed age-related sensory reductions. In one study [54], more than half of the AD patients erred on a color-matching test, most often in the blue category (see also [58]), as opposed to only 10% of healthy age-matched older adults. More recently, deficiencies in color discrimination for AD patients were found to be unrelated to the severity of AD [59] or to neuropsychological assessment [60], suggesting that "such deficits are not secondary consequences of attentional or motivational impairment" ([61], p. 2580; cf., [54]). These AD-related color-vision deficits may be linked to specific

lesions in the brain caused by the disease on top of visual sensory deficits that occur in healthy aging. For example, damage to the visual cortices of the brain, especially in peristriate and inferotemporal areas, due to AD-related increased neurofibrillary tangles in these areas [62], subcortical damage [63], or increase in retinal loss of ganglionar cells [64], may impair the ability to discriminate colors [55]. Similarly, there is evidence for AD-related damages to the lens (more severe than healthy age-related changes), related to aggregation of AD-related proteins [65], further potentially limiting color-vision.

Relating the evidence above to the Stroop test, Cronin-Golomb and colleagues [56] found that errors in the discrimination of font colors (specifically, blue) could account for some of the variance in color-naming latencies for AD patients. Moderate evidence for an increase in Stroop effects related to reduced color discrimination in AD is given by Fisher and colleagues [25] (we note that the authors only reported a trend), leading to the conclusion that there is a need for "... exercising caution in interpreting Stroop color-word test scores in demented older adults. Further investigation is needed to determine the extent of color-vision deficits in AD patients ... " (p. 757). In contrast to color-naming, reading aloud is found to be relatively preserved in AD as compared to age-matched controls [66]. Reading aloud is a highly practiced process, [67, 68] and less affected by aging or AD. Therefore, in the current analysis, we will examine whether the expected large impact AD has on color-naming versus its minor impact on reading leads to a significant increase in DI; and if this inflates DI, in turn, can it point to a sensory source of an AD-related increase in Stroop effects.

The current study

The goal of this study was to conduct a critical update on the literature on Stroop and AD, presenting a meta-analysis of 18 studies published since 1990 with 1,641 AD and healthy aged participants. To foreshadow our results, in the first analysis, we will show a significant effect for AD on Stroop effects, SI and on Ci, beyond age-related effects documented in the literature [5]. In the second analysis, we will demonstrate the limited, yet significant, role of AD-related slowing in speed of processing (as compared to healthy older adults) in explaining deteriorated Stroop performance in AD. The third analysis examines a subset of seven studies that presented data for an additional reading (Rn) block. We will show a very strong link between AD-related changes in color-naming and in reading,

supporting a general increase in dimensional imbalance (DI) for AD patients. Finally, a comparison of the relative contributions of cognitive slowing and DI to the AD-related increase in Stroop effects will demonstrate that both are comparable in explanatory power, although DI showed a marginally higher contribution.

METHODS

Selection of sample studies

The studies included in this analysis, dating back to 1990, were collected (July–September 2011) by accessing several electronic databases (i.e., Medline, PsycInfo, PubMed, Scopus, and Google Scholar), using a diverse list of search terms and key words (see Supplementary Material). The references in the retrieved articles were also consulted in order to identify additional suitable studies. We selected only studies that included: (a) data for two groups of older adults (over 65 years old), namely, a group of AD patients, as well as a group of healthy, non-demented controls, matched for age and gender; (b) measures of Cn, representing latencies for naming aloud the color of colored patches (or strings of colored X's or colored dots) or the font-color of neutral words, and of Ci, representing latencies for naming aloud the font color of incongruent color-words (e.g., the word "BLUE" printed in red ink); and (c) words which were presented in a Latin alphabet[2].

Given that there is no single accepted universal behavioral criterion to diagnose AD [69], we have based our inclusion criteria on three commonly used instruments to assess the presence of AD: (a) criteria from the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA), which require confirmation of cognitive impairment by neuropsychological testing (and occasionally post-mortem histological evidence) in order to accurately classify possible, probable or definite AD [70]. From the range of classifications provided in this assessment, we include studies for which the participants were rated as having at least probable AD; (b) the Clinical Dementia Rating (CDR) scale [71], which is used to assess cognitive performance in six specific areas: orientation, judgment and problem solving,

---

[2]We excluded studies that used a character-based text (e.g., Chinese) rather than an alphabetic form. The literature suggests differential effects for these two linguistic systems, in terms of underlying phonology [111], brain activation [112], and Stroop performance [113].

memory, home and hobbies, community affairs, and personal care. The inclusion criterion for studies investigating AD groups based on this scale was a score of equal to or greater than 0.5 (=very mild) on a range of 0 = none, 0.5 = very mild, 1.0 = mild, 2.0 = moderate, to 3.0 = severe; and (c) the Mattis Dementia Rating Scale (MDRS), a 36-task and 32-stimulus card instrument administered to an individual, where a score of less than 130 was used for inclusion as it indicates some level of AD [72]. It is important to note that in most studies two or more criteria were used to ensure the presence of AD[3].

Of the 87 Stroop with AD studies found in our literature search, 18 met the above depicted criteria. The main reasons for excluding studies were lack of Cn or Ci tasks or lack of a control group of healthy individuals. A list of the studies included, along with their respective characteristics, can be found in Table 1. The table also includes the level of AD severity for the AD groups in each study. A subset of seven studies included measures for Rn, which represents latencies for reading aloud color-words printed in black ink (e.g., the word "RED" printed in black ink, on a white background); these are indicated as such in Table 1.

Data pooling

In each of the 18 studies, we included one average score representing individuals with AD and another single score representing healthy matched controls. These groups comprise a total of 1,641 participants: 637 AD patients, with an average age of 75.1 (SD = 4.9) years, and 977 controls, with an average age of 73.5 (SD = 4.0) years. For the analysis of DI, we selected only the seven studies that provided Rn data, and these comprise a total of 394 participants: 208 AD patients and 186 age-matched controls. In studies that included two or more AD groups [23, 73], the data in Table 1 represents a weighted average of these groups; in other cases we included only the patient group that met our criteria (possible AD, [74]; AD, [35, 75–76][4] . In three

studies [76–78] the data were read off a bar figure (Fig. 1, p. 750, Fig. 1, p. 364, and Fig. 1, p. 306, respectively), and in two other studies [79, 80] the original data were received via personal communications with the authors[5] .

The data for each study and for each task (as presented in Table 1) were averaged to present the mean response time in ms per item. For six studies that used the Golden (or Golden-like[6]) version of the Stroop task [25, 35, 73, 79, 81, 82], averages were obtained by dividing the time limit (45 s or 90 s) by the number of responses acquired within this period. In these cases, the standard deviations were estimated by multiplying the latter average values by the original coefficients of variation. For eight other studies [29, 75, 76, 80, 83–86], which used renditions of the Victoria[7] version of the Stroop task, averages and standard deviations were obtained by dividing the time to complete the task by the number of items presented.

Data analysis

To compare latencies between AD and control groups across studies, we conducted paired-sample *t*-tests, based on the assumption of a normal distribution in such large samples. The speed of processing account was tested by conducting a Brinley [41] analysis, plotting Cn and Rn latencies for AD patients as a function of healthy age-matched controls across the 18 studies. DI was tested with data from seven studies that provided values for Rn, Cn, and Ci latencies.

To verify effect sizes, accounting for the variance of the means within each study, the number of participants, and the possibility that studies that did not find significant effects were not published, the analysis of group differences was complemented by meta-analysis statistics (with a syntax found in [87], and functions provided in [88]), based on a subset of 15 studies that provided standard deviations[8]. Meta-analysis statistics

---

[3]Many researchers classify "very mild" Alzheimer's (e.g., based on a score of 0.5 on the CDR scale) as mild cognitive impairment (MCI [114–116], see also [117]). The examined studies that include very mild AD specify that "recruitment and assessment methods permit the diagnosis of DAT [=Dementia of Alzheimer's Type] in individuals who elsewhere may be characterized as MCI" [78] (p. 748, the terminology was added in brackets). Notably, recent studies have suggested more extensive criteria allowing a better classification of patients into the various subtypes of MCI [118, 119].

[4]In several studies we included groups of AD patients that matched our criteria, excluding groups of patients diagnosed with vascular dementia [35], with FTD [75, 76], or whose neuropsy-

chological assessment scores significantly deviated from the overall mean of the sample [81] (based on MDRS scores). From the study by Spieler and colleagues [23], we included the control group entitled "young-old," whose mean age matched the overall mean of the sample.

[5]Thus, the data may diverge from the reported means in these papers. For example, in [79], we report the average of 16 patients and a matched group of 16 control participants.

[6]In the Golden [120] version of the Stroop test, performance is gauged as the number of items correctly completed in 45 s.

[7]In the Victoria [121] version of the Stroop test, performance is gauged as the time it takes to name the color of 24 items (Cn or Ci).

[8]Meta-analyses can also be performed on studies that report *t*-test values or correlation values, even in the absence of standard

Table 1
Data extracted from the 18 Alzheimer's and control groups, as used in this meta-analysis

Study	Severity	Data Persons with AD					Controls					Derivatives Dementia		Control	
		n	Age	Ci	Cn	Rn	n	Age	Ci	Cn	Rn	SI	DI	SI	DI
Amieva, 2002 [83]	Mild[a,b]	28	75.8	1191	626	–	28	75.2	974	644	–	565	–	330	–
Balota, 2010 [109]	Poss-Severe[a,b]	12	91.8	1224	1038	–	35	85.0	1134	964	–	186	–	170	–
Bayard, 2011 [84]	Mild[a]	40	77	2083	662	–	40	77.5	1829	738	–	1421	–	1091	–
Belleville, 2006 [82]	Mild/Mod[a,b]	12	72.5	4029	1525	730	12	72.7	1632	782	477	2505	795	850	305
Berardi, 2005 [81]	Mild[a]	5	66	1552	776	489	20	67	1154	692	464	774	287	462	228
Bondi, 2002 [73]	Poss/Mil/Mod[a,b]	47	72.4	3303	1192	692	51	73.4	1563	838	570	2111	500	725	268
Collette, 2007 [75]	Mild[a,b]	25	72.5	1425	934	–	28	70.6	985	719	–	491	–	266	–
Coubard, 2011 [79]	Mild/Mod[a]	17	78.68	2946	1035	662	17	77.65	2243	824	589	1911	373	1419	235
Duchek, 2009 [78]	Possible[a,d]	71	71.2	1200	910	–	220	75.2	900	810	–	290	–	90	–
Fisher, 1990 [25]	Moderate[c]	36	70.1	4639	1376	692	36	72.9	1347	693	466	3263	684	654	227
Hogge, 2008 [85]	Mild[a,b]	17	76.1	1376	851	–	20	74.4	1082	742	–	525	–	340	–
Joubert, 2010 [86]	Mild[a]	16	74.2	2750	867	–	16	72.4	1600	710	–	1883	–	890	–
Levinoff, 2004 [29]	Mild[a]	30	74.1	2717	1261	–	23	73.0	1316	680	–	1456	–	636	–
McGuinness, 2010 [35]	Mild/Mod[a]	75	77.7	2616	1034	615	28	70.2	1522	733	471	1582	419	789	262
Murtha, 2002 [76]	Mild/Mod[e]	8	75.4	1967	958	–	10	77.3	1208	729	–	1009	–	479	–
Spieler, 1996 [23]	Poss/Mild[a,d]	62	73.4	1690	1409	–	25	70.5	1069	894	–	281	–	175	–
Tse, 2010 [77]	Possible[a,b]	74	75.8	1200	950	–	22	71.8	950	810	–	250	–	140	–
Zarei, 2010 [80]	Poss/mild[a,b]	16	79.5	1803	847	547	22	70.7	1219	608	462	956	300	611	146

Note: Data represent mean naming latencies in milliseconds per item, see text for details. Poss-possible; Mil-mild; Mod-moderate. The following criteria were used to assess the severity of Alzheimer's disease: [a]National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) Criteria; [b]Mattis Dementia Rating Scale (MDRS); [c]Blessed Dementia Scale (BDS); [d]Clinical Dementia Rating (CDR) Scale; [e]Foldstein Mini-Mental Status Exam (MMSE).

provide tools to summarize data and observe significant trends across various studies, and calculate an effect size, thus providing a strong and meaningful indicator for effects that occur in the population [88]. Similarly, when calculating regressions, we have included adjusted r^2 statistics (adj-r^2) that take into account the effect of the sample size. All tests were performed with an alpha level set at 0.05.

RESULTS

AD-related differences in SI

In the right-most panel of Table 1, we present the data on SI. Observe that in all studies, the SI for the AD groups was larger than for the control group. Indeed, the average SI for all 18 AD groups was significantly larger than the average value for control groups (1191 versus 558 ms, $t(17) = 4.02$, $p = 0.001$). This difference was confirmed in a meta-analysis with 15 studies. Homogeneity was rejected for both Cn and Ci ($\chi^2 > 20.3$, $p < 0.01$, for both), hinting at the possibility of external sources of variance between studies. Thus, we conducted a mixed model unstandardized effect size test[9] of the AD-related differences in Cn and Ci across the 15 studies [88]. For both tasks we found effect sizes to be significantly different from zero, indicating a significant AD-related increase. The effect size for the AD-related increase in Ci ($M = 711.3$, SD = 118.2) was significantly larger than for the AD-related increase in Cn ($M = 247.2$, SD = 53.9; $t(28) = 13.8$, $p < 0.001$). Since SI is the difference in latency between Ci and Cn, this confirms the significance of the increase in SI for the AD groups. In sum, both a simple latency comparison and a meta-analysis modeling support a significant AD-related increase in SI as compared to healthy adults, across studies and labs.

Testing the speed of processing hypothesis

Figure 1 (A, B) presents Brinley analyses of latencies for Ci (circles) and Cn (triangles), across all 18 examined studies, plotting the scores of the AD groups as a function of the scores on the same variables for control groups. We note that data from one study, Fisher and colleagues [25], indicated by unfilled markers in Fig. 1, diverges from the groups' mean (their ratio of latencies for AD group as a function of con-
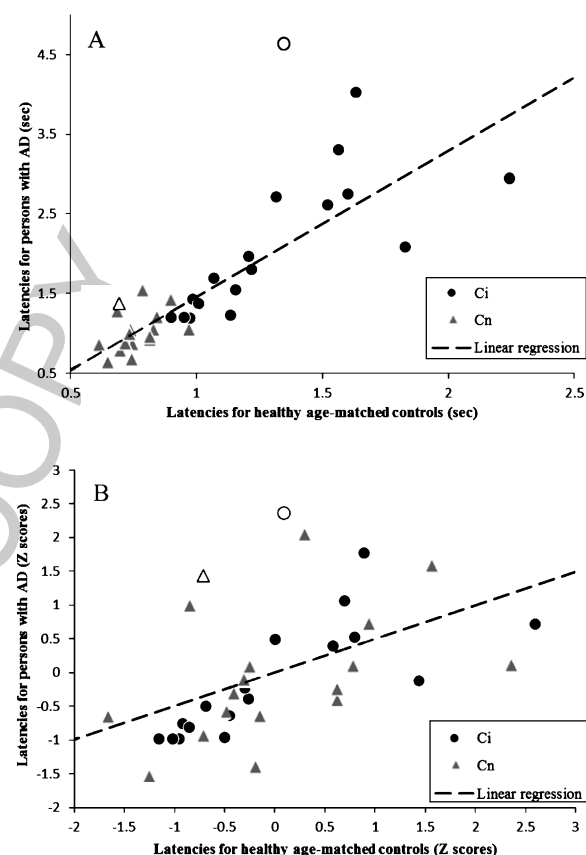

Fig. 1. Brinley plots for color-naming neutral stimuli (Cn, triangles) and color-naming incongruent words (Ci, circles) for 18 groups of persons with AD as a function of similar latencies from 18 groups of age-matched controls. Unfilled symbols represent data from Fisher and colleagues [25], which was omitted from a follow-up regression; broken lines represent the linear regression for 34 data points (Cn and Ci). Panel A presents a plot for raw-latency data, and Panel B presents a plot for normalized data with z scores calculated for Ci and Cn for AD and control groups separately. See text for more details.

trols was more than two standard deviations away from the mean ratio for both tasks across all studies). As a consequence, the following analyses exclude data from this study, leaving 17 studies. Note the parallel analysis for the complete set of 18 studies is presented in the Supplementary Material, showing very similar results.

Traditional Brinley analysis

Figure 1a shows the traditional Brinley analysis for the latency data. To investigate whether a single function is sufficient to describe AD-related differences in both Ci and Cn [6, 11], we fitted the following regression equation to the data[9]:

---

deviations of the mean. However, we wanted to conform to a standard analysis on all studies.

[9]We note that in the meta-analysis by Amieva and colleagues [15] a non-linear regression was used to conduct a Brinley analysis.

$$\begin{cases} Cn_{AD} = a_n + b_n * Cn_{control} \\ \\ Ci_{AD} = a_i + b_i * Ci_{control} \end{cases} \quad (4)$$

This four-parameter model was compared with a limited two-parameter model, where the same slope ($b_n = b_i$) and intercept ($a_n = a_i$) were used for both conditions. Increasing the number of parameters from 2 to 4 did not result in a significant reduction in variance ($F(2,28) = 1.01$, $p > 0.1$). Therefore, a single linear regression ($r^2 = 0.74$, $adj\text{-}r^2 = 0.73$, $F(1,32) = 94.7$, $p < 0.001$), with a significant slope value of 1.8 ($t(32) = 9.7$, $p < 0.001$) and an intercept that was not significantly different from zero ($t(32) = 1.79$, $p > 0.05$), is sufficient to describe the data for both tasks.

Note that in an SOP framework, each of the individual linear regressions for Cn and for Ci (line 1 and 2 in Eq. 4, respectively) are expected to significantly explain a large portion of their respective variance. Yet, the separate regression line for Cn was only marginally significant ($F(1,15) = 4.3$, $p = 0.06$) accounting for less than 20% of the variance in the data ($adj\text{-}r^2 = 0.17$). The large extent of the explained variance for the combined model might be thus the outcome of the biased distribution of latencies, with Ci latencies about twice as large as Cn latencies. This biased distribution increased the weight of the Ci data points in the regression disproportionally.

Normalized Brinley analysis

To control for the above noted bias in the traditional Brinley analysis, we normalized latency data across the 17 studies for Ci and Cn task in each age group, generating four new Z-score data sets (after confirming that all four original distributions, AD-Ci, control-Ci, AD-Cn, and control-Cn, fit a normal curve based on a Kolmogorov-Smirnov test). This transformation fits our purposes well because: a) It preserves the separate regression functions for Ci and for Cn (explained variance and slopes, as the intercepts were not significantly different from zero); and b) it equates the mean ($=0$) and standard deviation ($=1$) of both distributions, thus also equating their weight in a Brinley analysis. Next, we replicated the Brinley analysis on the normalized data as presented in Fig. 1b. As in the traditional analysis, using two rather than one single regression function did not result in a significant

reduction in variance ($F(2,28) < 1.0$). However, the single regression line can explain only about one third of the variance due to AD in the normalized data ($r^2 = 0.38$, $adj\text{-}r^2 = 0.36$, $F(1,32) = 19.4$, $p < 0.001$; with a significant slope of 0.6, $t(32) = 4.4$, $p < 0.001$ and an intercept of zero).

In sum, with a traditional Brinley analysis, the speed of processing model can explain close to 75% of the variance due to AD. After correcting for a bias in the distribution of the Ci versus Cn data, the explained variance dropped to 36%. These explained variances dropped even further if the analyses were to re-include a single study [25] that varied significantly from the mean (61% and 22% for the original and normalized data, respectively; see Supplementary Material). Furthermore, we note that a speed of processing model could only marginally explain the variance due to AD in Cn latencies (in either the normalized or the original data).

Testing the color-vision (DI) hypothesis

Group comparison

As shown in Table 1, within the seven studies that provided such information, the DI was larger for the AD groups than for their matched control groups (averages of 480 versus 239 ms, $t(6) = 3.8$, $p < 0.01$). In a meta-analysis, using a mixed model of unstandardized[10] effect sizes (homogeneity was rejected, $\chi^2 > 100$, $p < 0.001$ for both tasks) we found the AD-related differences in Rn and in Cn across the seven studies to be significantly different from zero. The systematic increase in DI (the difference between Cn and Rn) for the AD groups was confirmed by the fact that the effect size for the AD-related increase in Cn ($M = 320.2$, $SD = 63.6$) was significantly larger than for the AD-related increase in Rn ($M = 128.0$, $SD = 27.4$; $t(12) = 7.3$, $p < 0.001$). In sum, an AD-related increase in DI was confirmed in both a simple and unstandardized model.

AD-related differences in Cn and Rn

Figure 2 plots AD-related differences in Rn as a function of AD-related differences in Cn, across seven studies. Group differences on Rn were highly predictable by group differences on Cn ($r^2 = 0.981$, $adj\text{-}r^2 = 0.963$, $F(1,5) = 128.4$, $p < 0.001$), with a significant slope of 0.33 ($t(5) = 11.3$, $p < 0.001$), and an intercept that was not significantly different from zero ($t(5) < 1$).

In the current data-set, we followed the more commonly used Brinley linear analysis, minimizing the number of free parameters.

[10]An unstandardized effect size test is a test conducted on the raw data (unstandardized means). Raw data is used to accumulate effects that are always measured on the same scale (e.g., ms).
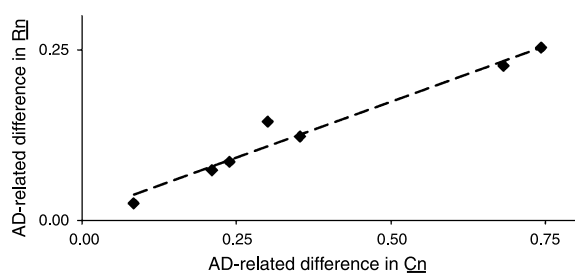
Fig. 2. Differences in RT between AD and age-matched control groups, for Rn (reading color-neutral words) as a function of Cn (color-naming neutral stimuli) in seven Stroop and AD studies. The broken line represents the linear regression.

Based on this analysis, the AD-related increase in Cn is closely coupled to the increase in Rn, but is three-fold larger in its extent.

Comparing the SOP with the color-naming (DI) hypothesis

Which interpretation for AD-related differences in Stroop effects can explain a larger portion of the variance, DI or SOP? The analyses so far provide some limited support for SOP and more substantial support for DI. To distinguish the two, we compared their relative contribution in explaining AD-related changes in Stroop. Stroop effects were measured using Ci latencies, since SI (Ci-Cn) is linearly related to both DI (Cn-Rn) and Cn. For the same reasons (and to maximize the number of studies that can be used for this analysis), Cn was used as an estimate of SOP [7, 11]. The regression analysis on changes in Ci attributed to AD ($Ci_{AD} - Ci_{Control}$) as a function of changes in DI ($DI_{AD} - DI_{Control}$) can explain 85% of the variance ($r^2 = 0.87$, $adj\text{-}r^2 = 0.85$, $F(1,5) = 34.3$, $p = 0.002$, with the seven studies that provide the relevant data). The regression analysis on AD-related changes in Ci as a function of changes in Cn can explain 64% of the variance ($r^2 = 0.67$, $adj\text{-}r^2 = 0.64$, $F(1,15) = 30.1$, $p < 0.001$, using 17 studies)[11]. These correlations indicate that contributions of DI and Cn (representing SOP) toward Ci are considerable, with a slightly higher contribution resulting from DI (85% versus 64%. The difference, however, was not significant, $z < 1$). We note that this apparent difference is effaced in a lim-

ited comparison of the power of DI and SOP, using only the seven studies that include latencies for reading (Cn and Ci: $r^2 = 0.87$, $adj\text{-}r^2 = 0.84$, $F(1,5) = 32.9$, $p < 0.005$).

In sum, it appears that one cannot refute the contribution of either factor (i.e., degraded color-naming and cognitive slowing) to changes in Stroop performance following AD. Nonetheless, the impact of SOP on AD-related changes in Stroop performance, above and beyond those found in healthy older individuals, appears to be modest (explaining only a third of the variance). Furthermore, the findings indicate that the contribution of sensory degradation on Stroop performance following AD, again above and beyond what is found in healthy aging, may be larger than previously estimated.

DISCUSSION

In a meta-analysis of 18 Stroop and AD studies, published across two decades (1990–2011), we found conclusive evidence for an AD-related increase in Stroop effects. This effect has generally been taken as a reflection of a severe reduction in selective attention following AD. However, we also found evidence for two alternative sources that can, in part, explain these AD-related differences, viz., a slowdown of speed of processing and color-vision degradation. The current paper presents a new comprehensive approach for examining how cognitive performance is evaluated in AD patients, using the Stroop test, the most common test for selective attention as an example. Specifically, we used a larger number of studies than available for a previous meta-analysis on this topic (e.g., 7 studies were used by Amieva et al. [15]), tested a novel hypothesis (sensory degradation, [6, 10, 11]), and used a sophisticated mathematical model for latency. The results of these analyses led to four main conclusions presented in more detail in the next sections. First, AD-related deterioration of selective attention processes may not be as severe as previously suggested. Second, the Stroop test provides a biased measure for selective attention in the AD population, as it is confounded by other factors. Third, sensory degradation in AD plays a significant role in tests of cognitive performance, and should be addressed. Finally, based on these findings we suggest measures found in the literature that can be taken to control for the speed and sensory biases in measures of selective attention (and other cognitive functions) in AD.

---

[11]This regression does not include the data from Fisher and colleagues [25] similar to the Brinley analysis. After adding this study to the regression, including all 18 studies, the percent of explained variance is slightly increased ($r^2 = 0.70$, $adj\text{-}r^2 = 0.68$, $F(1,16) = 37.3$, $p < 0.001$).

Speed of processing (SOP)

In a traditional Brinley analysis, we found that plotting latencies (for color-naming neutral and incongruent trials, Cn and Ci) for AD groups as a function of control groups, a generalized slowing in SOP could explain around 75% of the variance. The slope of the function indicated that latencies for AD patients were approximately 80% slower than latencies for age-matched controls across both tasks. This value corresponds to an approximate estimated increase of 65% in latencies based on a Brinley analysis by Nebes and Brady [24], who examined latencies for various cognitive tasks that did not include Stroop. These results question the traditional selective attention hypothesis on the Stroop test. In selective attention, AD-related changes in Stroop effects reflect differences in attentional processes. Thus, the proportion of AD-related changes in Ci (a task that demands selective attention), must be significantly larger than the proportion of AD-related changes in Cn (a task that demands no selective attention). In contrast, our analysis indicated that a single function, relating latencies for both Ci and Cn from AD-patients to latencies for age-matched controls, was sufficient to account for a significant portion of the variance.

Our Brinley analysis results are to some extent at odds with the results from a meta-analysis on AD and Stroop by Amieva and colleagues [15], which did not find conclusive support for SOP as a mediating factor. The stronger impact of SOP in our data may relate to two main differences between the two meta-analyses. First, the Amieva meta-analysis investigated data from only seven published studies. In the current analysis, we were able to almost triple the number of studies to 18 (or 17 in a restrictive analysis), as several new papers were published in the past decade. With more data to work with, our analysis had more power to detect such differences. Second, in the Amieva analysis, the data from three different control groups were replicated to create a 10-group control data set used for comparison with 10 different AD groups (to increase the number of data points, they included seven separate AD-groups from three studies). As a consequence, the variance in control-group data was artificially reduced. To avoid this problem, we averaged data across AD groups from a particular study, thus including only one set of data points (Ci, Cn) for AD patients and another single set for matched controls for each individual study.

A further investigation of our Brinley analysis results revealed a discrepancy; the separate function relating Cn latencies for AD and control groups was only marginally significant and explained less than 20% of the variance. However, the analysis of both tasks combined explained nearly 75% of the joint variance. This was likely the result of a bias in the distribution of latencies, where Ci values were twice as large as those of Cn. To address this issue, we added a Brinley analysis on normalized data points, preserving the original Ci and Cn distributions, while equating their averages and standard deviations. In the normalized analysis, an SOP model was again significant, but could only account for about one third of the joint variance. Taking both Brinley analyses together, our results show that an AD-related generalized slowing (in addition to age-related slowing) can explain some of AD-related differences in Stroop performance (even though it is not clear to what extent). In other words, the general finding on an AD-related increase in Stroop effects may not simply represent a large decrease in selective attention. Rather, it reflects, in part, a change in the speed of processing related to AD. Following our results, it seems a prudent strategy to account for SOP in analyzing Stroop data from AD patients. This can be done by co-varying external measures of speed of processing that are part of the assessment battery for AD patients. For example, one might assess SOP using the Montreal Cognitive Assessment (MoCA, [89]) or the Severe Impairment Battery (SIB, [90, 91]). An alternative route is provided by Spieler and colleagues [23] who divided Stroop effects by latencies in baseline conditions (Cn and/or Rn) controlling thus for the impact of SOP on Stroop. Other researchers have used somewhat different methods to control for these variables. For example, Bayard and colleagues [84] used the ratio between latencies to name the color of color words and naming the color of colored dots. Bondi et al. [73] used yet another ratio [(C - CW)/C] to control for color naming performance associated with processing a color stimulus, for lexical access of color words, and for the final motor pathways activated by the articulation of color–word responses.

Sensory degradation

In a separate analysis of a subset of seven AD and Stroop studies, we found evidence for a disproportionate AD-related slowing in baseline color-naming versus only a moderate AD-related slowing in reading. Comparing AD patients to healthy older adults, we found that AD-related changes in Cn were three times as large as changes in Rn. The relationship between these variables can explain as much as 98% of the variance, showcasing the universal nature of an

increase in the DI measure (the difference between color-naming and reading speed) across studies and labs. The data corresponds to a meta-analysis by Ben-David and Schneider [6] on healthy aging, finding that age-related changes in Cn were twice as large as in Rn. Thus, older adults with AD have a unique deficit in color-naming in the Stroop context that exceeds healthy age-related changes. As mentioned in the outset to this paper, this may be related to the specific neuropathology seen in AD impacting on their color vision [54–57], while largely sparing reading performance. As a result, it is reasonable to assume that in the Ci task, AD patients have to inhibit the lexical content for a longer time than their age-matched counterparts until the color information is processed (see also [6, 10, 11]). Indeed, this increased dimensional imbalance accounted for the majority (85%) of the AD-related variance in Stroop performance (Ci), echoing Fisher's early warning [25] to be cautious in interpreting AD patients' results on the color-word Stroop task.

The link between sensory and cognitive abilities has been widely documented with an aging population on general cognitive performance (see a seminal paper [92]; note recent work relating cognitive performance in aging to hearing [93, 94], and to vision [95]), and specifically with respect to selective attention [96]. Taken together, the apparent decline in performance with aging may arise because the information delivered by the sensory system becomes degraded, as suggested by the information degradation framework [3, 10, 97]. Since sensory decline is intensified by AD, it is reasonable to assume that a further decline in performance ensues. In particular for the Stroop task, degraded color input, due to a decline in color-vision with AD, can hinder performance.

The color naming task (Cn) involves identification, thus, impaired lexical access is a plausible alternative explanation for the elevated latencies for color-naming in AD. Key problem with this explanation is that lexico-semantic impairment will also affect access to the meaning of the color-words and as a consequence should attenuate the Stroop interference. However, Duong et al. [98] found that intentional access is impaired earlier in the course of AD than automatic access. Reading being an automatic, well-practiced function may be less impaired than color naming. More support for a sensory explanation can be found in Salmone's et al. [99] study, using a Color discrimination test (Farnsworth-Munsell 100 hue test) that is minimally dependent on cognition, as it does not directly require recognition of colors. In their study, Salmone and colleagues found that color discrimina-

tion performance worsens with severity of AD. AD patients' error scores on the color discrimination were significantly higher than the scores of matched controls even after controlling for MMSE scores. These findings led the authors to suggest that "Color discrimination impairment in AD is not a purely cognitive problem but seems to be related to damage of the structures responsible for sensation and/or perception of color stimuli" (p. 505).

The results of our analysis receive further support from literature on AD and object naming, where AD patients' performance was found to be impaired to a larger extent when they were asked to identify an object depicted in color images, rather than in monochrome (specifically for identification of living things; see a meta-analysis [100], but see [101], for opposite results in a change-detection task). Additional support can be found in Rizzo et al.'s study [102]. They found no significant difference between participants with AD and controls on measures of visual acuity and stereo acuity but found a significant difference in color vision. Furthermore, Rizzo et al. found that impairment of color vision correlates with overall severity of cognitive impairment in AD.

Future studies are needed to investigate the specific role of visual sensory decline (in color-vision and in general) on Stroop performance and on other cognitive tests following AD. Furthermore, it is notable that in aging, and specifically in AD, there is a large variability in the sensory and cognitive abilities of individuals [103]. It is therefore strongly advisable that in administering color-naming Stroop (or any other cognitive test), one should also assess the sensory acuity of patients. Testing patients' color vision as well as their visual acuity is of great importance as most neuropsychological tests are administered in the visual modality. Clinicians and researchers interested in measuring selective attention in AD should also consider using alternative tools that do not involve color-vision. As an example, take the 'number Stroop' task, which uses two printed Arabic numerals and requires individuals to indicate the larger one, based on either the numerical value or font size [104]. Other options include a subset of the Attention Network Test ([105]; for an example with AD patients, see [106, 107]), and the Flanker test [108]. Recently, a new task was suggested to improve the sensitivity of the Stroop task for predicting cognitive decline—'Stroop switching task', a combination of the Stroop task and the 'task-switching paradigm' [109, 110]. In a study by Hutchison et al. [110], incongruent error rates from the Stroop switching task discriminated healthy aging from AD patients

better than any of the other 18 cognitive tasks used in the study.

Caveat

It is important to acknowledge other factors that can affect Stroop performance in AD patients. The level of dementia severity may account for at least some of the variance. However, the number of studies included was not sufficient to conduct such an analysis (see also [15]). Another factor that may contribute to the variance observed is the colors used in the studies included in the analyses. Deterioration of color vision in AD patients is disproportionate. For example, Cronin-Golomb and colleagues [56] found more errors in the discrimination of blue font color for AD patients. As data on colors used and latencies for specific color was not consistently provided in the examined studies, we were not able to test for this aspect. Finally, in our introduction, we cite papers to support the effect of reduced color discrimination in AD on Stroop performance. We note that not all of these papers present significant effects (e.g., Fisher and colleagues [25] reported only a trend).

Summary

In a meta-analysis and modeling of 18 studies on Stroop with AD patients, we show a significant increase in Stroop effects for AD patients, as compared to healthy older adults. Our analysis does not provide support for a unique link between an increase in Stroop effects and a decrease in selective attention associated with AD; rather, this increase appears to be mediated by degraded color vision, a slowing in SOP, or a combination of these two factors. We do not postulate that there are no AD-related changes in selective attention. However, our analysis indicates that the impact of AD on selective attention, above and beyond age-related changes, may not be as large as it was estimated to be previously, using the color-word Stroop task. Based on these findings, we suggest complementing the Stroop task (in a neuro-psychological assessment of AD patients) by measures of sensory acuity, SOP, and alternative measures for selective attention that better control for these factors. In conclusion, the current results highlight the impact of sensory degradation on behavioral measures of cognitive performance in aging and neuro-pathological populations, and advise caution in the interpretation of the findings from cognitive tests in these populations, when sensory and SOP factors have not been accounted for.

SUPPLEMENTARY MATERIAL

Supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-131244.

REFERENCES

[1] Hasher L, Zacks R (1988) Working memory, comprehension, and aging: A review and a new view. *Psychol Learn Motiv Adv Res Theory* **22**, 193-225.

[2] Hasher L, Zacks R, May C (1999) Inhibitory control, circadian arousal, and age. *Attn Perform* **17**, 653-675.

[3] Schneider BA, Pichora-Fuller MK, Daneman M, In S, Gordon-Salant RD, Frisina AN (2010) Effects of senescent changes in audition and cognition on spoken language comprehension. In *The Aging Auditory System*, Gordon-Salant RD, Frisina AN, Popper AN, Fay RR, eds. Springer, New York, pp. 167-210.

[4] Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* **18**, 643-662.

[5] McDowd J, Shaw R (2000) Attention and aging: A functional perspective. In *The Handbook of Aging and Cognition*, Craik FIM, Salthouse TA, eds. Lawrence Earlbaum Associates, New Jersey, pp. 221-292.

[6] Ben-David BM, Schneider BA (2009) A sensory origin for color-word Stroop effects in aging: A meta-analysis. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* **16**, 505-534.

[7] Trenerry MR, Crosson B, DeBoe J, Leber WR (1989) *Stroop neuropsychological screening test manual*. Psychological Assessment Resources, Odessa, FL.

[8] Bagiella E, Novack T, Ansel B, Diaz-Arrastia R, Dikmen S, Hart T, Temkin N (2010) Measuring outcome in traumatic brain injury treatment trials: Recommendations from the traumatic brain injury clinical trials network. *J Head Trauma Rehab* **25**, 375-382.

[9] Shim YS, Kim JS, Shon YM, Chung YA, Ahn KJ, Yang DW (2008) A serial study of regional cerebral blood flow deficits in patients with left anterior thalamic infarction: Anatomical and neuropsychological correlates. *J Neurol Sci* **266**, 84-91.

[10] Ben-David BM, Schneider BA (2010) A sensory origin for color-word Stroop effects in aging: Simulating age-related changes in color-vision mimics age-related changes in Stroop. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* **17**, 730-746.

[11] Ben-David BM, Nguyen LLT, van Lieshout PHHM (2011) Stroop effects in persons with traumatic brain injury: Selective attention, speed of processing, or color-naming? A meta-analysis. *J Int Neuropsychol Soc* **17**, 354-363.

[12] Troyer AK, Leach L, Strauss E (2006) Aging and response inhibition: Normative data for the Victoria Stroop Test. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* **13**, 20-35.

[13] West R, Alain C (2000) Effects of task context and fluctuations of attention on neural activity supporting performance of the Stroop task. *Brain Res* **873**, 102-111.

[14] Verhaeghen P, Cerella J (2002) Aging, executive control, and attention: A review of meta-analyses. *Neurosci Biobehav Rev* **26**, 849-857.

[15] Amieva H, Phillips LH, Della Sala S, Henry JD (2004) Inhibitory functioning in Alzheimer's disease. *Brain* **127**, 949-964.

[16] World Health Organization (1992) The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines.

[17] Lyketsos CG, Olin J (2002) Depression in Alzheimer's disease: Overview and treatment. *Biol Psychiatry* **52**, 243-252.

[18] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM (2007) Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement* **3**, 186-191.

[19] Burns A, Jacoby R, Luthert P, Levy R (1990) Cause of death in Alzheimer's disease. *Age Ageing* **19**, 341-344.

[20] Beard CM, Kokmen E, Sigler C, Smith GE, Petterson T, O'Brien PC (1996) Cause of death in Alzheimer's disease. *Ann Epidemiol* **6**, 195-200.

[21] Crols R, Saerens J, Noppe M, Lowenthal A (1986) Increased GFAp levels in CSF as a marker of organicity in patients with Alzheimer's disease and other types of irreversible chronic organic brain syndrome. *J Neurol* **233**, 157-160.

[22] Bettner LG, Jarvik LF, Blum JE (1971) Stroop color-word test, non-psychotic organic brain syndrome, and chromosome loss in aged twins. *J Gerontol* **26**, 458-469.

[23] Spieler DH, Balota DA, Faust ME (1996) Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *J Exp Psychol Hum Percept Perform* **22**, 461-479.

[24] Nebes RD, Brady CB (1992) Generalized cognitive slowing and severity of dementia in Alzheimer's disease: Implications for the interpretation of response-time data. *J Clin Exp Neuropsychol* **14**, 317-326.

[25] Fisher LM, Freed DM, Corkin S (1990) Stroop Color-Word Test performance in patients with Alzheimer's disease. *J Clin Exp Neuropsychol* **12**, 745-758.

[26] MacLeod CM (1992) The Stroop task: The 'gold standard' of attentional measures. *J Exp Psychol Gen* **121**, 12-14.

[27] Barkley RA (2001) The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychol Rev* **11**, 1-29.

[28] Parasuraman R, Haxby J (1993) Attention and brain function in Alzheimer's disease: A review. *Neuropsychology* **7**, 242-272.

[29] Levinoff EJ, Li KZH, Murtha S, Chertkow H (2004) Selective attention impairments in Alzheimer's disease: Evidence for dissociable components. *Neuropsychology* **18**, 580-588.

[30] Perry R, Hodges J (1999) Attention and executive deficits in Alzheimer's disease: A review. *Brain* **122**, 383-404.

[31] Baudic S, DallaBarba G, Thibaudet M, Smagghe A, Remy P, Traykov L (2006) Executive function deficits in early Alzheimer's disease and their relations with episodic memory. *Arch Clin Neuropsychol* **21**, 15-21.

[32] Perry RJ, Watson P, Hodges JR (2000) The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: Relationship to episodic and semantic memory impairment. *Neuropsychologia* **38**, 252-271.

[33] Posner MI, Petersen SE (1990) The attention system of the human brain. *Annu Rev Neurosci* **13**, 25-42.

[34] Smith EE, Jonides J (1999) Storage and executive processes in the frontal lobes. *Science* **283**, 1657-1661.

[35] McGuinness B, Barrett SL, Craig D, Lawson J, Passmore AP (2010) Attention deficits in Alzheimer's disease and vascular dementia. *J Neurol Neurosurg Psychiatry* **81**, 157-159.

[36] Birren J (1965) Age changes in speed of behavior. In *Behaviour, Aging, and the Nervous System*, Welford TA, Birren JE eds. Charles C. Thomas, Springfield, pp. 191-216.

[37] Salthouse TA (1985) Speed of behavior and its implications for cognition. *Handbook of the Psychology of Aging*, Birren JE, Schaie KW, eds. Van Nostrand Reinhold, New York, pp. 400-426.

[38] Cerella J (1985) Information processing rates in the elderly. *Psychol Bull* **98**, 67-83.

[39] Cerella J (1990) Aging and information processing rate. In *Handbook of the Psychology of Aging*, Birren JE, Schaie KW, eds. Academic Press, San Diego, pp. 201-221.

[40] Verhaeghen P (2011) Aging and executive control: Reports of a demise greatly exaggerated. *Curr Dir Psychol Sci* **20**, 174-180.

[41] Brinley J (1965) Cognitive sets, speed and accuracy of performance in the elderly. In *Behavior Aging and the Nervous System*, Welford TA, Birren JE, eds. C. C. Thomas, Springfield, pp. 114-149.

[42] Thapar A, Ratcliff R, McKoon G (2003) A diffusion model analysis of the effects of aging on letter discrimination. *Psychol Aging* **18**, 415-429.

[43] Nebes RD, Madden DJ (1988) Different patterns of cognitive slowing produced by Alzheimer's disease and normal aging. *Psychol Aging* **3**, 102-104.

[44] Verhaeghen P, De Meersman L (1998) Aging and the Stroop effect: A meta-analysis. *Psychology Aging* **13**, 120.

[45] Foster J, Behrmann M, Stuss D (1995) Aging and visual search: Generalized cognitive slowing or selective deficits in attention? *Aging Cogn* **2**, 279-299.

[46] Werner JS, Steele VG (1988) Sensitivity of human foveal color mechanisms throughout the life span. *J Opt Soc Am A* **5**, 2122-2130.

[47] Nguyen-Tri D, Overbury O, Faubert J (2003) The role of lenticular senescence in age-related color vision changes. *Invest Ophthalmol Vis Sci* **44**, 3698-3704.

[48] Salthouse TA, Meinz EJ (1995) Aging, inhibition, working memory, and speed. *J Gerontol B Psychol Sci Soc Sci* **50**, P297-P306.

[49] Akutsu H, Legge GE, Ross JA, Schuebel KJ (1991) Psychophysics of reading–X. Effects of age-related changes in vision. *J Gerontol* **46**, P325-P331.

[50] Van der Elst W, Van Boxtel M, Van Breukelen G, Jolles J (2006) The Stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* **13**, 62-79.

[51] Melara RD, Algom D (2003) Driven by information: A tectonic theory of Stroop effects. *Psychol Rev* **110**, 422-471.

[52] Sabri M, Melara RD, Algom D (2001) A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *J Exp Psychol Hum Percept Perform* **27**, 515-537.

[53] Eidels A, Townsend J, Algom D (2010) Comparing perception of Stroop stimuli in focused versus divided attention

paradigms: Evidence for dramatic processing differences. *Cognition* **114**, 129-150.

[54] Cronin-Golomb A, Corkin S, Rizzo JF, Cohen J, Growdon JH, Banks KS (1991) Visual dysfunction in Alzheimer's disease: Relation to normal aging. *Ann Neurol* **29**, 41-52.

[55] Cronin-Golomb A, Sugiura R, Corkin S, Growdon JH (1993) Incomplete achromatopsia in Alzheimer's disease. *Neurobiol Aging* **14**, 471-477.

[56] Cronin-Golomb A, Corkin S, Growdon JH (1995) Visual dysfunction predicts cognitive deficits in Alzheimer's disease. *Optom Vis Sci* **72**, 168-176.

[57] Mendola JD, Cronin-Golomb A, Corkin S, Growdon JH (1995) Prevalence of visual deficits in Alzheimer's disease. *Optom Vis Sci* **72**, 155-167.

[58] Cohen J, Cronin-Golomb A, Growdon J, Corkin S (1988) Color vision deficits in Alzheimer's disease. *Soc Neurosci Abst* **219**, 14.

[59] Pache M, Smeets CHW, Gasio PF, Savaskan E, Flammer J, Wirz-Justice A, Kaiser HJ (2003) Colour vision deficiencies in Alzheimer's disease. *Age Ageing* **32**, 422-426.

[60] Salamone G, Di Lorenzo C, Mosti S, Lupo F, Cravello L, Palmer K, Musicco M, Caltagirone C (2009) Color discrimination performance in patients with Alzheimer's disease. *Dement Geriatr Cogn Disord* **27**, 501-507.

[61] Adlington RL, Laws KR, Gale TM (2009) Visual processing in Alzheimer's disease: Surface detail and colour fail to aid object identification. *Neuropsychologia* **47**, 2574-2583.

[62] Braak H, Braak E, Kalus P (1989) Alzheimer's disease: Areal and laminar pathology in the occipital isocortex. *Acta Neuropathol* **77**, 494-506.

[63] Sadun AA, Bassi CJ (1990) Optic nerve damage in Alzheimer's disease. *Ophthalmology* **97**, 9-17.

[64] Blanks JC, Torigoe Y, Hinton DR, Blanks RH (1996) Retinal pathology in Alzheimer's disease. I. Ganglion cell loss in foveal/parafoveal retina. *Neurobiol Aging* **17**, 377-384.

[65] Goldstein LE, Muffat JA, Cherny RA, Moir RD, Ericsson MH, Huang X, Bush AI (2003) Cytosolic β-amyloid deposition and supranuclear cataracts in lenses from people with Alzheimer's disease. *Lancet* **361**, 1258-1265.

[66] Friedman RB, Ferguson S, Robinson S, Sunderland T (1992) Dissociation of mechanisms of reading in Alzheimer's disease. *Brain Lang* **43**, 400-413.

[67] Besner D, Stolz JA, Boutilier C (1997) The Stroop effect and the myth of automaticity. *Psychon Bull Rev* **4**, 221-225.

[68] Mansfield JS, Legge GE, Bane MC (1996) Psychophysics of reading. XV: Font effects in normal and low vision. *Invest Ophthalmol Vis Sci* **37**, 1492-1501.

[69] McKeel DW, Price JL, Miller JP, Grant EA, Xiong C, Berg L, Morris JC (2004) Neuropathologic criteria for diagnosing Alzheimer disease in persons with pure dementia of Alzheimer type. *J Neuropathol Exp Neurol* **63**, 1028-1037.

[70] Blacker D, Albert MS, Bassett SS, Go RC, Harrell LE, Folstein MF (1994) Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease. The National Institute of Mental Health Genetics Initiative. *Arch Neurol* **51**, 1198-1204.

[71] Morris J (1997) Clinical Dementia Rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr* **9**, 173-176.

[72] Monsch AU, Bondi MW, Salmon DP, Butters N, Thal LJ, Hansen LA, Wiederholt WC, Cahn DA, Klauber MR (1995) Clinical validity of the Mattis Dementia Rating Scale in detecting Dementia of the Alzheimer type. A double cross-validation and application to a community-dwelling sample. *Arch Neurol* **52**, 899-904.

[73] Bondi MW, Serody AB, Chan AS, Eberson-Shumate SC, Delis DC, Hansen LA, Salmon DP (2002) Cognitive and neuropathologic correlates of Stroop Color-Word Test performance in Alzheimer's disease. *Neuropsychology* **16**, 335-343.

[74] Berardi AM, Parasuraman R, Haxby JV (2005) Sustained attention in mild Alzheimer's disease. *Dev Neuropsychol* **28**, 507-537.

[75] Collette F, Amieva H, Adam S, Hogge M, Van der Linden M, Fabrigoule C, Salmon E (2007) Comparison of inhibitory functioning in mild Alzheimer's disease and frontotemporal dementia. *Cortex* **43**, 866-874.

[76] Murtha S, Cismaru R, Waechter R, Chertkow H (2002) Increased variability accompanies frontal lobe damage in dementia. *J Int Neuropsychol Soc* **8**, 360-372.

[77] Tse CS, Balota DA, Yap MJ, Duchek JM, McCabe DP (2010) Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology* **24**, 300-315.

[78] Duchek JM, Balota DA, Tse CS, Holtzman DM, Fagan AM, Goate AM (2009) The utility of intraindividual variability in selective attention tasks as an early marker for Alzheimer's disease. *Neuropsychology* **23**, 746-758.

[79] Coubard OA, Ferrufino L, Boura M, Gripon A, Renaud M, Bherer L (2011) Attentional control in normal aging and Alzheimer's disease. *Neuropsychology* **25**, 353-367.

[80] Zarei M, Patenaude B, Damoiseaux J, Morgese C, Smith S, Matthews PM, Barkhof F, Rombouts S, Sanz-Arigita E, Jenkinson M (2010) Combining shape and connectivity analysis: An MRI study of thalamic degeneration in Alzheimer's disease. *Neuroimage* **49**, 1-8.

[81] Berardi AM, Parasuraman R, Haxby JV (2005) Sustained attention in mild Alzheimer's disease. *Dev Neuropsychol* **28**, 507-537.

[82] Belleville S, Rouleau N, Van der Linden M (2006) Use of the Hayling task to measure inhibition of prepotent responses in normal aging and Alzheimer's disease. *Brain Cogn* **62**, 113-119.

[83] Amieva H, Lafont S, Auriacombe S, Le Carret N, Dartigues JF, Orgogozo JM, Fabrigoule C (2002) Inhibitory breakdown and dementia of the Alzheimer type: A general phenomenon? *J Clin Exp Neuropsychol* **24**, 503-516.

[84] Bayard S, Erkes J, Moroni C, Collège des Psychologues Cliniciens spécialisés en Neuropsychologie du Languedoc Roussillon (CPCN, Languedoc Roussillon) (2011) Victoria Stroop Test: Normative data in a sample group of older people and the study of their clinical applications in the assessment of inhibition in Alzheimer's disease. *Arch Clin Neuropsychol* **26**, 653-661.

[85] Hogge M, Salmon E, Collette F (2008) Interference and negative priming in normal aging and in mild Alzheimer's disease. *Psychol Belg* **48**, 1-23.

[86] Joubert S, Brambati SM, Ansado J, Barbeau EJ, Felician O, Didic M, Lacombe J, Goldstein R, Chayer C, Kergoat MJ (2010) The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia* **48**, 978-988.

[87] Alferes V (2003) Meta-analysis: Fixed and random effects models, http://www.spsstools.net/, Accessed on January 15, 2012.

[88] Lipsey M, Wilson D (2001) *Practical Meta-Analysis*. Sage Publications Inc, USA.

[89] Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H

(2005) The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* **53**, 695-699.

[90] Schmitt FA, Ashford W, Ernesto C, Saxton J, Schneider LS, Clark CM, Ferris SH, Mackell JA, Schafer K, Thal LJ (1997) The severe impairment battery: Concurrent validity and the assessment of longitudinal change in Alzheimer's disease. *Alzheimer Dis Assoc Disord* **11**(Suppl 2), S51-S56.

[91] Panisset M, Roudier M, Saxton J, Boller F (1994) Severe impairment battery: A neuropsychological test for severely demented patients. *Arch Neurol* **51**, 41-45.

[92] Lindenberger U, Baltes PB (1994) Sensory functioning and intelligence in old age: A strong connection. *Psychol Aging* **9**, 339-355.

[93] Lin F, Ferrucci L, Metter E, An Y, Zonderman A, Resnick S (2011) Hearing loss and cognition in the Baltimore longitudinal study of aging. *Neuropsychology* **25**, 763-770.

[94] Ben-David BM, Chambers CG, Daneman M, Pichora-Fuller MK, Reingold EM, Schneider BA (2011) Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *J Speech Lang Hear Res* **54**, 243.

[95] Henderson R, Allerhand M, Patton N, Pattie A, Gow A, Dhillon B, Starr JM, Deary IJ (2011) Vision and intelligence at age 83 in the Lothian Birth Cohort 1921. *Intelligence* **39**, 148-154.

[96] Guerreiro MJS, Murphy DR, Van Gerven PWM (2010) The role of sensory modality in age-related distraction: A critical review and a renewed view. *Psychol Bull* **136**, 975-1022.

[97] Schneider B, Pichora-Fuller M (2000) Implications of perceptual deterioration for cognitive aging research. *The Handbook of Aging and Cognition.* Lawrence Erlbaum Associates Publishers, USA, pp. 155-219.

[98] Duong A, Whitehead V, Hanratty K, Chertkow H (2006) The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia* **44**, 1928-1935.

[99] Salamone G, Di Lorenzo C, Mosti S, Lupo F, Cravello L, Palmer K, Caltagirone C (2009) Color discrimination performance in patients with Alzheimer's disease. *Dement Geriatr Cogn Disord* **27**, 501-507.

[100] Laws K, Adlington R, Gale T, Moreno-Martínez F, Sartori G (2007) A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia* **45**, 2674-2682.

[101] Neargarder SA, Cronin-Golomb A (2005) Characteristics of visual target influence detection of change in naturalistic scenes in Alzheimer disease. *Cogn Behav Neurol* **18**, 151-158.

[102] Rizzo M, Anderson SW, Dawsonc J, Nawrotd M (2000) Vision and cognition in Alzheimer's disease. *Neuropsychologia* **38**, 1169.

[103] Martin A, Brouwers P, Lalonde F, Cox C, Foster NL, Chase TN (1986) Towards a behavioral typology of Alzheimer's patients. *J Clin Exp Neuropsychol* **8**, 594-610.

[104] Girelli L, Sandrini M, Cappa S, Butterworth B (2001) Number-Stroop performance in normal aging and Alzheimer's-type dementia. *Brain Cogn* **46**, 144-149.

[105] Fan J, McCandliss BD, Sommer T, Raz A, Posner MI (2002) Testing the efficiency and independence of attentional networks. *J Cogn Neurosci* **14**, 340-347.

[106] Fernandez-Duque D, Black SE (2006) Attentional networks in normal aging and Alzheimer's disease. *Neuropsychology* **20**, 133-143.

[107] Fuentes LJ, Fernández PJ, Campoy G, Antequera MM, García-Sevilla J, Antúnez C (2010) Attention network functioning in patients with dementia with Lewy bodies and Alzheimer's disease. *Dement Geriatr Cogn Disord* **29**, 139-145.

[108] Collette F, Schmodt C, Scherrer C, Adam S, Salmon E (2009) Specificity of inhibitory deficits in normal aging and Alzheimer's disease. *Neurobiol Aging* **30**, 875-889.

[109] Balota DA, Tse CS, Hutchison KA, Spieler DH, Duchek JM, Morris JC (2010) Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in Stroop color naming. *Psychol Aging* **25**, 208-218.

[110] Hutchison KA, Balota DA, Duchek JM (2010) The utility of Stroop task switching as a marker for early stage Alzheimer's Disease. *Psychol Aging* **25**, 545-559.

[111] Spinks JA, Liu Y, Perfetti CA, Tan LH (2000) Reading Chinese characters for meaning: The role of phonological information. *Cognition* **76**, B1-B11.

[112] Qiu J, Luo Y, Wang Q, Zhang F, Zhang Q (2006) Brain mechanism of Stroop interference effect in Chinese characters. *Brain Res* **1072**, 186-193.

[113] Biederman I, Tsao YC (1979) On processing Chinese ideographs and English words: Some implications from Stroop-Test results. *Cogn Psychol* **11**, 125-132.

[114] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment: Clinical characterization and outcome. *Arch Neurol* **56**, 303-308.

[115] Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B (2001) Current concepts in mild cognitive impairment. *Arch Neurol* **58**, 1985-1992.

[116] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.

[117] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, Cummings JL, de Leon M, Feldman H, Ganguli M, Hampel H, Scheltens P, Tierney MC, Whitehouse P, Winblad B (2006) Mild cognitive impairment. *Lancet* **367**, 1262-1270.

[118] Petersen RC, Negash S (2008) Mild cognitive impairment: An overview. *CNS Spectr* **13**, 45.

[119] Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Jack Jr CR (2009) Mild cognitive impairment: Ten years later. *Arch Neurol* **66**, 1447.

[120] Golden C (1978) *The Stroop Color and Word Test.* Stoelting Company, Chicago, IL.

[121] Regard M (1981) Stroop test: Victoria version. Canada University of Victoria, Department of Psychology, doi: 101207s15326942dn2801_2