APPLIED NEUROPSYCHOLOGY: ADULT
https://doi.org/10.1080/23279095.2017.1373281

Virtual Apartment-Based Stroop for assessing distractor inhibition in healthy aging

Thomas D. Parsons and Michael Barnett

Department of Psychology, University of North Texas, Denton, Texas, USA

ABSTRACT
Executive functioning declines found in healthy aging may impact negatively on everyday activities. Inhibitory control is an executive function that involves the preclusion of irrelevant information and prohibition of prepotent responses. Decreased inhibitory control has been suggested as a feature of age-related cognitive decline. In this study, we compared a Virtual Apartment Stroop Task with paper-and-pencil and computerized Stroop modalities in older adults and young adults. The primary results were (a) within groups the classic "Stroop pattern" found in traditional modalities (e.g., paper-and-pencil; computer automated) was observed in the Virtual Apartment Stroop Task; (b) the Virtual Apartment "Stroop scores" between groups were consistent with significant differences found between older adults and young adults on the paper-and-pencil and computerized Stroop tasks; and (c) individuals in the older adult cohort performed significantly more poorly on the Virtual Apartment Stroop Task than did participants in the young adult cohort. These results suggest the potential of the Virtual Apartment Stroop Task to distinguish between prepotent response inhibition and resistance to distractor inhibition in aging adults.

KEYWORDS
Aging; ecological validity; executive functioning; neuropsychology; Stroop; virtual reality

Introduction

Inhibitory control involves the preclusion of irrelevant information and prohibition of prepotent responses (Hasher, Lustig, & Zacks, 2007). Decreased inhibitory control has been suggested as a feature of age-related cognitive decline (Andres & Van der Linden, 2000; Hasher et al., 2007; Kim, Rasher, & Zacks, 2007). Furthermore, deficiencies in inhibitory control have been proposed as prominent early impairments in Alzheimer's disease (Bélanger, Belleville, & Gauthier, 2010; Belleville, Rouleau, & Van der Linden, 2006; Collette, Schmidt, Scherrer, Adam, & Salmon, 2007), which suggests that such deficits may develop during preclinical stages of the disease.

The supervisory control of attention to pertinent features of the environment while ignoring irrelevant, though salient, information is an important aspect of goal-directed behavior (Baddeley, 1996; Baddeley & Hitch, 1974; Norman & Shallice, 1986). Supervisory attentional control is typically measured by putting task-relevant information in conflict with task-irrelevant information. An example can be found in the classic Stroop Color Naming Task (Stroop, 1935) that presents color word stimuli (e.g., "blue") printed in colored font (e.g., red) to participants who have been instructed to specify the font color while ignoring the printed word. Scoring of inhibitory control typically involves assessment of differences in accuracy and reaction time between congruent (the word "red" is in red font color) and incongruent (the word "red" is in blue font color). Deficits in inhibitory control have been associated with age-related cognitive declines in healthy aging (Andrés, Guerrini, Phillips, & Perfect, 2008; Aschenbrenner & Balota, 2017; Clapp, Rubens, Sabharwal, & Gazzaley, 2011; Coxon et al., 2014; Healey, Campbell, & Hasher, 2008). The Stroop Task has been used to examine the cumulative magnitude of the interference found with increases in age (Aschenbrenner & Balota, 2015; Bugg, DeLosh, Davalos, & Davis, 2007; Jackson & Balota, 2013; Puccioni & Vallesi, 2012).

Multi-item presentation of Stroop stimuli

Clinical neuropsychologists often use a paper-and-pencil version of the Stroop task to assess executive functioning and inhibitory control through the presentation of blocks of multiple Stroop stimuli on a card (multi-item presentation; Uttl & Graf, 1997). For example, the Delis–Kaplan Executive Function System (D-KEFS) Color–Word Interference Test (Delis,

CONTACT Thomas D. Parsons thomas.parsons@unt.edu Department of Psychology, University of North Texas, 1155 Union Circle #311280, Denton, TX 76203, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hapn.

Kaplan, Kramer, & Ober, 1997) is a multi-item paper-and-pencil Stroop, in which participants are presented with a card that has 50 incongruent color words. Although multi-item presentations of rows of Stroop stimuli have been validated in multiple studies, single-item presentation of Stroop stimuli may be preferable in situations where neuropsychologists are interested in: precise analysis of reaction times for individual stimuli. Moreover, multi-item presentations may result in confounds from or interactions with visual distractor interference. Older adults have been found to have greater sensitivity to visual distraction found in multiple-item presentations of cognitive stimuli (Lustig, Tonev, & Hasher, 2000). These results suggest that the differences between the multiple-item and single item presentations may interact with age (see Davidson, Zacks, & Williams, 2003). A single-item presentation may be preferable when assessing older adults.

Single-item presentation of Stroop stimuli

A number of computer automated assessments of single-item Stroop-like tasks have been developed: Automated Neuropsychological Assessment Metrics (ANAM) Stroop (Johnson, Vincent, Johnson, Gilliland, & Schlegel, 2008); CogSport card recognition task (CogSport, 1999); HeadMinder picture recognition (Erlanger, Feldman, & Kutner, 1999); and Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT Color Match; Lovell, Collins, Podell, Powell, & Maroon, 2000). Of these computerized approaches, the ANAM single-item Stroop Task is perhaps closest to traditional multi-item Stroop presentations (see Reeves, Winter, Bleiberg, & Lane, 2007). Declines in executive functions such as inhibition impair neurocognitive and daily functioning (Aretouli & Brandt, 2010). Although both multi-item paper-and-pencil and single-item computerized versions of the Stroop have been used widely, their lack of everyday distractors may diminish predictions about real-world functioning (Chan, Shum, Toulopoulou, & Chen, 2008).

Everyday distractors impacting inhibitory control

Traditional Stroop measures do not reflect the various environmental distractions that occur in everyday inhibitory control. The relations between an older adult's performance on a Stroop test and behavior in various situations with real-world distractors can be incomplete in that these traditional Stroop tests do not allow for assessment of real-world distractors on a participant's performance. The sorts of distractions encountered in everyday activities may encumber inhibitory control by increasing perceptual and cognitive load, which may limit the older-aged person's ability to perform important cognitive processes (Lavie, 2005). Inhibitory control across the lifespan may be more affected by distraction in persons of advancing age (Getzmann, Gajewski, & Falkenstein, 2013).

Virtual Reality Stroop Task

A Virtual Reality Stroop Task (VRST) has been developed that uses real-world auditory and visual distracters for assessment of inhibitory control in real-world environments (Parsons, Courtney, Arizmendi, & Dawson, 2011; Parsons, Courtney, & Dawson, 2013). Participants are immersed in a virtual high-mobility multipurpose wheeled vehicle as Stroop stimuli appear on the windshield. The original VRST was designed to measure supervisory attentional processing in conditions with and without both cognitive and affectively salient distractors. It emulated aspects of both paper-and-pencil (e.g., Stroop Color Word Interference) and ANAM versions of the Stroop task. Similar to the paper-and-pencil version of the Stroop, the VRST presents Stroop stimuli (though in single-item presentation) to assess simple attention, gross reading speed, divided attentional abilities, and executive functioning. Furthermore, the VRST, like the computerized Stroop (e.g., ANAM Stroop), automates the paper-and-pencil Stroop task and allows for assessment of reaction time to single-item presentations of Stroop stimuli.

Virtual Classroom Stroop Task

Leveraging the initial VRST paradigm, other virtual reality-based Stroop tasks have emerged. For example, a virtual classroom has been developed that present distractors in a virtual classroom environment while participants perform the Stroop Task. This virtual classroom Stroop paradigm allows for the assessment of external cognitive inhibition by requiring participants to resist distractions in the environment. The virtual classroom Stroop has been validated for use with healthy adolescents (Lalonde, Henry, Drouin-Germain, Nolin, & Beauchamp, 2013) and persons with autism (Parsons & Carlew, 2016).

Virtual Apartment Stroop Task

A further development of the Virtual Reality Stroop Task approach can be found in the Virtual Apartment-Based Stroop Task (Henry, Joyal, & Nolin, 2012). The virtual apartment superimposes Stroop stimuli on a television in the interior of a virtual apartment. A kitchen is located to the left of the television and a window is located to the right of the television.

Distracters occur in various locations throughout the virtual environment. Audio-visual distractors include: a school bus and other vehicles can be seen passing outside the living room window; a cell phone rings and vibrates on the coffee table (in front of participant); a toy robot moves and makes noise on the floor. Auditory distractors included: doorbell; clock alarm; vacuum cleaner; jack hammering outside; a person sneezing; and noise from a jet passing over the apartment. Visual distractors include: a woman walking around in the kitchen. Preliminary data from a small cohort of healthy adult participants found the Virtual Apartment-Based Stroop to be capable of eliciting the Stroop effect with bimodal stimuli. Moreover, the Virtual Apartment-Based Stroop correlated significantly with measures of the Elevator counting with distracters, the Continuous Performance Task (CPT-II), and the Stop-it Task (Henry et al., 2012). These preliminary findings suggest the potential of the Virtual Apartment-Based Stroop for measuring cognitive inhibition for healthy adults.

The current study aimed to assess the Virtual Apartment-Based Stroop for differentiating performance of older-aged participants from younger-aged participants. We compared performance of each group on three Stroop modalities (multi-item paper-and-pencil Stroop); single-item computerized Stroop (no distractions); and the virtual reality-based Stroop (with distractions). We aimed to assess whether the typical Stroop effect (decreased processing speed for incongruent stimuli) pattern found in the D-KEFS and ANAM occur in the Virtual Apartment-Based Stroop. The exploratory nature of the study and results is underscored.

Method

Participants

Two groups ($N = 89$) participated in a study designed to compare the performance of an older age cohort with a college age cohort on three Stroop modalities: (a) the paper-and-pencil Stroop Task (D-KEFS Color Word Interference Test; Delis, Kaplan, & Kramer, 2001; multi-item, unimodal); (b) computerized Stroop (ANAM; C-SHOP, 2007; single-item, unimodal); and (c) the Virtual Apartment-Based Stroop (single-item, bimodal Stroop stimuli in a simulated apartment). The order of assessments was counterbalanced across participants. All measures were administered by trained psychometrists who received specific training and supervision.

Thirty-nine community-dwelling older adults (74% female; mean age = 74.38; *SD* 8.70) took part in this study. Participants for the healthy older adult group were recruited via announcements in public and semipublic settings (e.g., health fairs, retirement communities, community centers). Older adult group participants were healthy community-dwelling volunteers with estimated full-scale IQ scores from the Wechsler Test of Adult Reading (Wechsler, 2001) in the average range (Mean = 106.34; $SD = 9.25$). They endorsed familiarity and comfort with computers. They rated their competency as somewhat experienced. Participants from the young adult group included 50 healthy college aged participants (86% female; mean age = 20.59; *SD* 3.18). Younger aged participants were recruited from the undergraduate population at a large university in the southwestern United States. Participants were tested individually and they received course credit for participation. Further, participants in the younger aged group had estimated full scale IQ scores from the Wechsler Test of Adult Reading in the average range (Mean = 106.38; $SD = 8.43$). They endorsed familiarity and comfort with computers. They rated their competency as experienced. Within and across groups, participants were comparable and no significant differences for age, sex, estimated full scale IQ, or computer comfort were noted.

Participants took part in a health history interview that gathered information related to their neurological and psychiatric history. Participant selection involved strict exclusion criteria to limit possible confounding effects of comorbid factors that may influence cognition: psychiatric (e.g., intellectual disability, psychotic disorders, diagnosed learning disabilities, attention deficit/hyperactivity disorder, or mood disorder). Moreover, participants were excluded if they endorsed substance-related disorders (within 2 years of evaluation) and/or known neurologic conditions (e.g., seizure disorders, closed head injuries with loss of consciousness greater than 15 minutes, and neoplastic diseases). No participants were excluded.

Procedures and materials

Upon arriving at the lab, participants were instructed in the study's procedure, risks and benefits, and alternative options (non-participation). Prior to participation, they signed written informed consent approved by the university's institutional review board. After informed consent was obtained, basic demographic information was gathered and participants responded to questions designed to measure computer experience and usage activities: how frequently participants use a computer (e.g., "How many hours per week do you spend on the computer?"); their perceived level of computer skill

on a Likert scale (1 – not at all to 5 – very skilled); e.g., "How many hours per week do you spend playing video games?"; and what type of games they play (e.g., role-playing, strategy, sports, etc.). All participants were administered the same battery of tests.

All participants completed informed consent before participating in the study. Testing was conducted in dedicated testing rooms. All participants were administered (counterbalanced) the same neuropsychological battery and were assessed in the Virtual Apartment.

Cognitive functioning

To make sure that the groups (older adult and younger adult) had comparable estimated full-scale intelligence scores, the Wechsler Test of Adult Reading (WTAR) was used (Wechsler, 2001). The WTAR offers estimation of pre-morbid intellectual and memory abilities in healthy older adults without advanced neurocognitive decline.

Delis Kaplan Executive Functioning System Color Word Interference Test

The D-KEFS (Color–Word Interference Test) was used for the multi-item paper-and-pencil presentation of the Stroop test. We followed the D-KEFS manual's prescribed approach to administration and the D-KEFS's "Scoring Assistant" software for scoring of the Color–Word Interference Test. The subject was seated at a desk and was presented with the following stimuli from the D-KEFS: (a) "color naming" card with 50 colored blocks—participants were instructed to state the colors as quickly as possible without skipping any or making mistakes; (b) "word reading" stimulus card with 50 color words printed in black ink: participants were instructed to read the words as quickly as possible without skipping any or making mistakes; (c) "color-word interference" card with 50 color names printed in a discrepant ink color: participants were instructed to name the color of the ink that the letters were printed in and not read the word; and (d) "color-word interference/switching" card, in which the participant performed the interference task if and only if the words (50 total words) did not have a box drawn around them: participants were instructed that if no box was drawn around the word, then they were to name the color of the ink and not read the words. However, if a word is inside a little box, the participant was to read the word and not name the ink color. Participants were timed for completion of each condition, and all incorrect responses were recorded. Per the D-KEFS manual, the examiner recorded the time taken by each participant to complete each stimulus card. Moreover, the examiner noted incorrect (including self-corrected) responses. A reaction time variable was created by dividing completion time by the number of stimuli completed (maximum 50). An accuracy variable was calculated by dividing the number of correct responses (not including self-corrected responses) by the total stimuli number of stimuli presented (50).

Computerized Stroop Task

The ANAM (single-item presentation) Stroop task requires the subject to press a computer key labeled red, green, or blue to identify each color stimulus presented. We followed the ANAM (2007) manual's approach to administration and scoring of the ANAM Stroop test. Each participant used the keyboard to take the ANAM Stroop test. A short practice session (10 stimuli) was provided just prior to the actual testing of each Stroop condition (i.e., color-naming, word-reading, and interferences tasks). There are three possible blocks of 50 trials for this test. In the first block, the words RED, GREEN, and BLUE are presented individually in black type on the display. The user is instructed to read each word aloud and to press a corresponding key for each word ("red" = 1; "green" = 2; and "blue" = 3). In the second block, a series of XXXXs is presented on the display in one of three colors (red XXXXs, green XXXXs, or blue XXXXs). The user is instructed to say the color of the XXXXs aloud and to press the corresponding key based on color. In the third block, a series of individual words ("RED," "GREEN," or "BLUE") are presented in a color that does not match the name of the color depicted by the word. The user is instructed to say the color of the word aloud rather than reading the actual word and to press the response key assigned to that color. The participants were asked to respond to each individual color stimulus as quickly as possible without making mistakes. Speed and duration of single-item stimulus presentations were participant defined, that is, that a new stimulus appeared only after the participant correctly identified the previous stimulus. The ANAM (2007) scores the Stroop as follows:

- Color-word score was calculated by (a) multiplying the number of correct colors named by the number of correct words named; and (b) dividing the product by the sum of number of correct colors named plus the number of correct words named.
- Interference score was calculated by taking the number correct on the interference task and subtracting the color-word score.

Virtual Apartment-Based Stroop

The Virtual Apartment-Based Stroop (ClinicaVR: Henry et al., 2012) is a construct-driven virtual

environment, in which Stroop stimuli are superimposed on a television in the interior of a virtual apartment (see Figure 1). Participants were seated in the living room, in front of a flat-screen television. A kitchen is located to the left of the television and a window is located to the right of the television. Two monitors were used: (a) one for displaying the launcher application, which is used by the examiner administering the test; and (b) another for displaying the participant's view of the virtual environment in the head-mounted display (HMD; eMagin Z800 with an InterSense InteriaCube 2+ attached for tracking). The HMD was used to recreate a 3D -like effect and the InterSense InteriaCube 2+ allowed participants to look 360° around themselves and explore the environment by turning their head. The task builds on the unimodal (visually mediated) Stroop and measures cognitive interference using reaction time, commission errors and omission errors, and reaction time variability. The Virtual Apartment-Based Stroop extends the traditional Stroop paradigm via the inclusion of bimodal (auditory and visually mediated) stimuli.

The Virtual Apartment-Based Stroop consists of two conditions: a block-based condition and (Color Naming) a word-based (Word Reading) condition. In the block-based conditions, a series of colored rectangles (red, blue, and green) appear on the flatscreen television within the environment while a female voice states the names of colors (red, blue, and green). The order of colors presented was consistent for each participant. Participants were instructed to click a mouse button as quickly as possible when the spoken color matches the color of the rectangle on the virtual flatscreen television, and to withhold a response if the colors do not match. A total of 144 stimuli were be presented, with 72 targets and 72 non-targets. Participants took part in a no-distraction (congruent: color naming and word reading) and distraction (incongruent–interference) condition for the block-based condition. While the typical congruent stimulus presentations were performed without distraction, the typical Stroop interference condition was accompanied by distractors. The duration of the block-based condition was 4.8 minutes with a 1000 millisecond inter-stimulus interval (ISI).

Figure 1. Virtual Apartment Stroop Task.

In the word-based condition, color words were presented on the virtual flatscreen television (red, blue, and green) in different ink colors (red, blue, and green). These stimuli were congruent (e.g., the word "blue" in blue ink) and incongruent (e.g., the word "blue" in red ink). The colors were stated in a female voice as in the block-based condition. Participants were instructed to click the mouse when the stated word matches the color of the word presented on the virtual flatscreen television, and to withhold a response if the stated word and presented ink color do not match. The word-based condition was designed to measure cognitive interference in addition the external interference control and motor inhibition assessed by the block-based condition. The duration of this condition was 4.8 minutes with a 1000 millisecond interstimulus interval (ISI). A total of 144 stimuli were presented, with 72 targets and 72 nontargets. The complete duration of the Virtual Apartment-Based Stroop was 19.2 minutes. Data collected will include: (a) mean reaction time, (b) correct responses, and (c) throughput (correct responses per minute of available response time; Thorne, 2006).

During the interference condition, distracters are added to the task. The distractor stimuli appear in different field of view locations in the environment. Some distracters are audio-visual: School Bus passing on the street and SUV viewed through window on the right; iPhone ringing and vibrating on the table (in front of participant); Toy Robot moving and making noise on the floor (center). Auditory distractors included: Crumple Paper (left); Drop Pencil (left); Doorbell (left); Clock (left) Vacuum Cleaner (right); Jack Hammer (right); Sneeze (left) Jet Noise (center). Visual distractors included: Paper airplane (flying from left to right in front of the participant), Woman walking in the kitchen (center). That condition was designed to assess reaction times (simple and complex), selective attention (matching the auditory and visual stimuli), and external interference control (environmental distracters).

Data analytics

All data were analyzed using SAS version 9.4 (Statistical Analysis System). Descriptive statistics were calculated for participant demographics and for results of the three

Stroop modalities: paper-and-pencil Stroop task; computerized Stroop; and Virtual Apartment-Based Stroop. An inspection of the pattern of missing data revealed that 3 (1 older adult and 2 younger adult) cases could be recovered by imputing only one missing value, and data from another older aged participant could be used after imputing two values. Missing data were imputed by either mean substitution or last case carried forward. For all analyses, the results with and without imputed data were comparable. Traditional behavioral data collected as dependent variables for Stroop tasks include accuracy (number of correct responses) and reaction time (e.g., latency of response). Comparison across stimulus modalities (paper-and-pencil: DKEFS; computerized: ANAM; and virtual reality: Virtual Apartment), a throughput algorithm was used due to its consideration of mean response time for correct responses (mean RT) and accuracy (percentage correct). Throughput units are reported as correct responses per minute of available response time (Thorne, 2006).

Data acquisition was performed using a MATLAB scoring program (Wu, Lance, & Parsons, 2013) to log specific events in the environment using time stamps with millisecond temporal accuracy. This scoring program was developed for the original VRST (Parsons et al., 2013) and was adapted specifically for this study to enable the assessment of performance validity (suboptimal effort) and screening for outliers to establish data integrity: (a) identification of outliers as observations exceeding three standard deviations from the median reaction time; and (b) exclusion of observations that are in both the top 1% in speed and simultaneously in the bottom 1% of accuracy. No outliers were identified in the current study results.

For all analyses of variance (ANOVAs) that included a variable with more than one degree of freedom, pairwise comparisons were used to identify the precise nature of any main effects. Additionally, all significant main effects and interactions were followed up with pairwise comparisons to determine the nature of these effects. Effect sizes were calculated for each analysis. Correlational analyses were used to assess for similarities across modalities. A sequentially rejective test procedure based on a modified Bonferroni inequality was used on significant results to prevent inflation of Type I error rates (Rom, 1990). Additionally, a Greenhouse–Geisser correction was used to assess for sphericity for all reported main effects and interactions. There were no instances that sphericity was violated.

Results

Groupwise comparison of performance on Virtual Apartment Stroop Task

For groupwise comparisons, a 2 (Age: Older and Younger) by 3 (Apartment Condition: Color, Word, Interference) analysis of variance was conducted to examine differences in overall performance (i.e., throughput among Stroop conditions; See Figure 2). There was a statistically significant interaction between the effects of Age and Apartment Condition on throughput, $F(1, 86) = 3.67$, $p = .02$. Furthermore, a

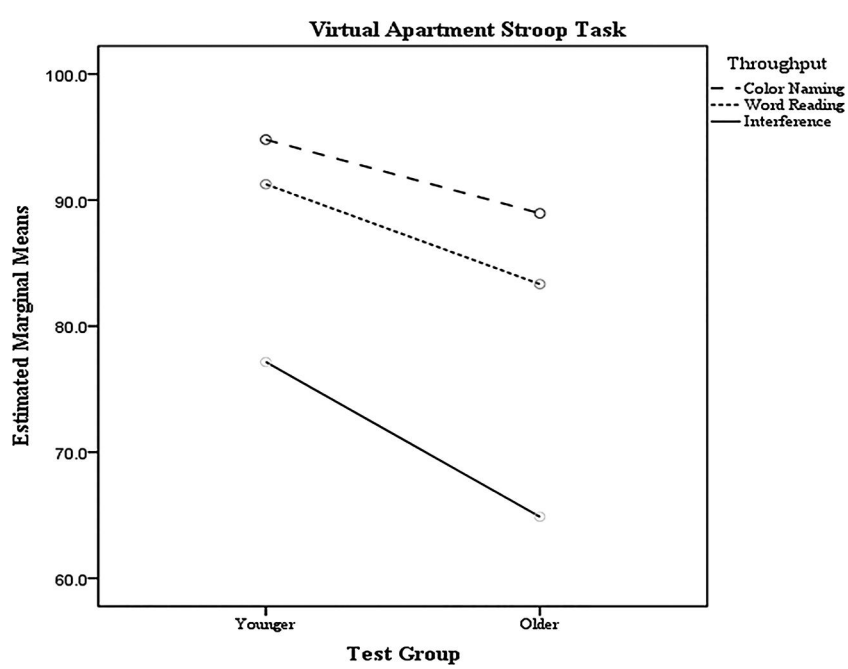

Figure 2. Comparison of conditions in Virtual Apartment Stroop Task throughput.

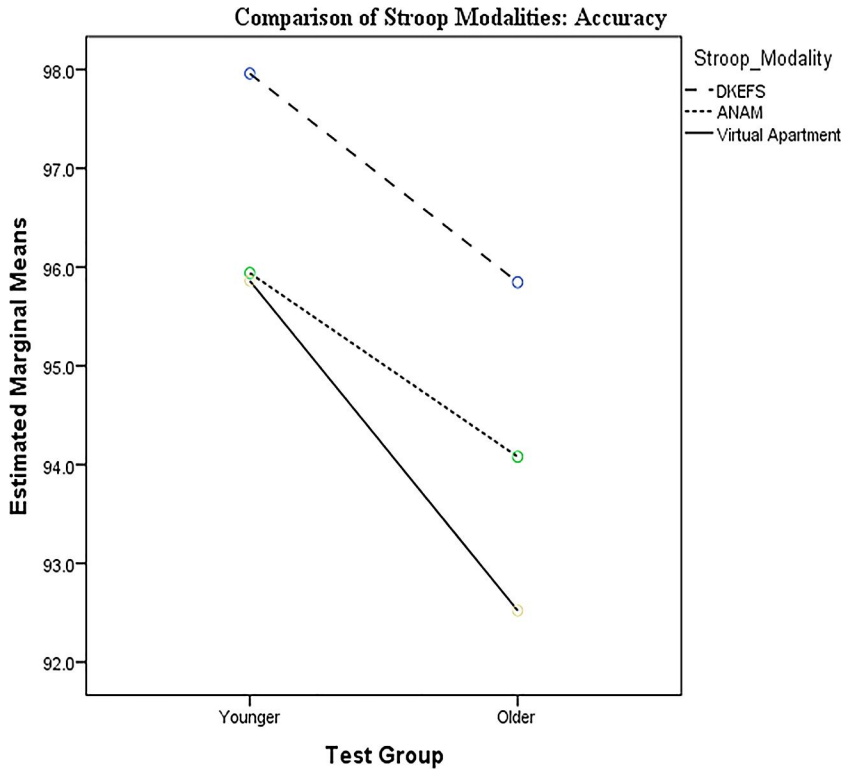

Figure 3. Comparison of Stroop modalities accuracy.

main effect for Age group was observed $F(1, 86) = 12.25$, $p < .001$, partial $\eta^2 = .13$. A main effect of Stroop conditions was also observed, $F(1, 86) = 163.67$, $p < .001$, partial $\eta^2 = .66$. Pairwise comparisons revealed significant differences (at the $p < .001$ level) between groups for color-naming (older adult $M = 88.94$, $SD = 12.21$; and young adult $M = 94.79$, $SD = 12.15$), word reading (older adult $M = 83.32$, $SD = 14.47$; and young adult $M = 91.25$, $SD = 14.10$), and interference (older adult $M = 64.87$, $SD = 15.11$; and young adult $M = 77.15$, $SD = 11.65$).

Comparison of multi-item presentation to single item presentations

For groupwise comparisons of accuracy, a 2 (Age: Older and Younger) by 3 (DKEFS; ANAM; Apartment modalities) analysis of variance was conducted to examine differences in accuracy among Stroop modalities (see Figure 3). While a significant interaction between the effects of Age and Stroop modality on accuracy were not observed, a main effect for Age group was present $F(1, 86) = 8.05$, $p < .005$, partial $\eta^2 = .09$. Furthermore, a main effect of Stroop conditions was observed, $F(1, 86) = 5.24$, $p < .006$, partial $\eta^2 = .66$. Pairwise comparisons revealed significant differences (at the $p < .001$ level) between groups for Virtual Apartment interference (older adult $M = 64.87$, $SD = 15.11$; and young adult $M = 71.71$, $SD = 11.65$), DKEFS Interference (older adult $M = 94.47$, $SD = 29.64$; and young adult $M = 54.75$, $SD = 12.73$), and ANAM interference (older adult $M = 34.49$, $SD = 12.95$; and young adult $M = 75.50$, $SD = 16.55$).

Comparison of the two single item presentations

To compare interference performance on the ANAM and Virtual Apartment between groups, a 2 (Age: Older and Younger) by 2 (ANAM interference throughput; Virtual Apartment interference throughput) analysis of variance was conducted. A significant interaction was observed between the effects of Age and Stroop modality on interference throughput $F(1, 86) = 55.92$, $p < .001$, A main effect for Age group was present $F(1, 86) = 126.57$, $p < .001$, partial $\eta^2 = .59$. Furthermore, a main effect of Stroop conditions was observed, $F(1, 86) = 69.50$, $p < .001$, partial $\eta^2 = .45$. Pairwise comparisons revealed significant differences (at the $p < .001$ level) between groups for Virtual Apartment interference throughput (older adult $M = 64.88$, $SD = 15.11$; and young adult $M = 77.56$, $SD = 11.66$) and ANAM interference throughput (older adult $M = 34.50$, $SD = 12.95$; and young adult $M = 75.51$, $SD = 16.56$).

Discussion

In this initial pilot study, we aimed to compare performance on the Virtual Apartment Stroop Task

(including distractors) to performance on paper-and-pencil (D-KEFS) and computerized (ANAM—no distractors) Stroop tasks. The primary results were (a) the classic "Stroop pattern" found in traditional modalities (e.g., paper-and-pencil; computer automated) was observed in the Virtual Apartment Stroop Task; (b) the Virtual Apartment "Stroop scores" were consistent with significant differences found between older adults and young adults on the paper-and-pencil and computerized Stroop tasks; and (c) individuals in the older adult cohort performed significantly more poorly on the Virtual Apartment Stroop Task (included distractors) than did participants in the young adult cohort. These results suggest that inhibiting prepotent responses in a virtual environment with distractors is more difficult for aging adults than inhibiting these responses in traditional assessments without distractors (e.g., ANAM Stroop). Specifically, in the Virtual Apartment, participants' responses were more accurate and had faster reaction times on tasks of inhibitory control. Furthermore, even greater differences were found between groups, with older adults having less accuracy and longer reaction times. Deficits in inhibitory control have been associated with age-related cognitive declines in healthy aging (Andrés et al., 2008; Aschenbrenner & Balota, 2017; Clapp et al., 2011; Coxon et al., 2014; Healey et al., 2008). The addition of everyday distractors using a virtual environment may be beneficial in situations where only subtle declines are present. Future work is needed to assess these preliminary findings.

Virtual Apartment Stroop elicited the classic "Stroop effect"

The Virtual Apartment Stroop Task was performed in a manner consistent with paper-and-pencil and computerized versions of the Stroop Task. In both groups, participants performed better (i.e., greater accuracy and faster response time) on color naming and word reading than they did on interference conditions. These findings are consistent with previous studies comparing traditional (paper-and-pencil and computerized) Stroop modalities to virtual reality versions of the Stroop (Armstrong et al., 2013; Henry et al., 2012; Lalonde et al., 2013; Parsons et al., 2013).

It is important to note that there are additional task demands when one moves from a paper-and-pencil assessment to a computer automated Stroop task. These differences are even more pronounced with the Virtual Apartment, in which participants had to match a visual stimulus to an auditory one (the virtual human naming a color) while environmental distractors called for greater attentional resources. As a result, the task demands are greater in the Virtual Apartment (including distractions), when compared to the classic Stroop Task (without distractors). Future work is needed to assess the extent to which these differences impact our understandings of inhibiting an overlearned response. Furthermore, there is need for future studies that offer a straightforward comparison between low and high distraction conditions within the virtual environment. A potential limitation of the current experimental design is that it relied on the computer automated Stroop that was completely devoid of environmental distractors for its non-distractor condition.

The current study is consistent with findings that have emphasized cognitive conflict processing that emerges between differing information sources (e.g., Stroop task). Given the greater levels of conflict processing in the presence of distractors, the Virtual Apartment findings may build on studies that have found that cognitive conflict processing can be impacted by the salience of a target stimulus (Kanske, 2012; Xue et al., 2013). Future work is needed to assess this possibility.

The decreased interference performance going from multi-item (paper-and-pencil) to single-item without distractors (computer automated), to single item with distractors (Virtual Apartment) may reflect the amount of time needed to read the word, which may potentially confound the Stroop interference scores. Chafetz and Matthews (2004) have argued that slower reaction times may be due to suppression of word reading in favor of color naming instead of a failure of response inhibition. However, the extent to which reading ability impacts the inhibition process is an open debate (Scarpina & Tagini, 2017). While our computer automated Stroop Task (ANAM Stroop) without distractors requires the participant to both press the button and state the color aloud, the Virtual Apartment Stroop Task (distractor condition) does not require the participant to read the word aloud. Furthermore, the findings of Chafetz and Matthews may reflect a multi-item presentation of stimuli and the fact that the words in their study had to be read aloud. Furthermore, the approach they used required logging by a human with notable limitations in response variance. Hence, the Virtual Apartment (single-item presentation, computer-automated logging, and lack of requirement for the participant to read aloud the response) may offer an alternative to older approaches. Moreover, the Virtual Apartment scoring approach uses a throughput algorithm that considers mean response time for correct responses (mean RT) and accuracy (percentage correct).

Practical implications of virtual Apartment Stroop findings

A number of practical implications can be inferred from the current study's findings. First, we demonstrated in that the Virtual Apartment Stroop Task successfully replicates the interference effect of classic Stroop tasks and may be used to assess cognitive interference control. While paper-and-pencil and computerized measures such as the DKEFS and the ANAM may provide adequate control, predictive value may be compromised by the lack of everyday distractors. These measures do not replicate the complex distractions that exist in the worlds in which individuals live and interact. Therefore, outcomes of this type of assessment may not accurately represent deficits in cognitive functioning. Virtual environment-based Stroop tasks may improve upon the shortcomings of paper-and-pencil and computer-based assessment by providing Stroop tasks in situations that include everyday distractors (Armstrong et al., 2013; Henry et al., 2012; Lalonde et al., 2013; Parsons & Carlew, 2016; Parsons et al., 2013). Importantly, this study highlights an important advantage of using virtual environments in psychological assessment: enhanced assessment using real-world-distractors. Rather than testing patients in sterile laboratory or clinical environments that do not reflect the complex distractors found in the real world, virtual environments allow for the patient to be assessed in a situation with the complex distractions that occur in the real world and the control of a laboratory.

Limitations and projections for future studies

Our findings should be understood in the context of some limitations. First, although the current study counterbalances order of presentation and distractor conditions, future studies are needed to assess the mediating effects of conflict processing differences between affective (including socially salient) and non-salient distractors. Researchers have recently extended the impact of emotionally salient stimulus on cognitive conflict processing to emotional conflict processing (Zinchenko, Kanske, Obermeier, Schröger, & Kotz, 2015). Adding socially and emotionally salient distractors to the Virtual Apartment platform may allow for more comprehensive deployment of participant resources. For example, the social and non-social distractors could be separated in future studies to include a high-social-distractor condition and neutral-distractor conditions. Furthermore, responses to emotionally salient distractors in the Virtual Apartment may enhance assessment of the complexities found in real-world interactions. Research has shown that conflict processing of affective information results in longer reaction times and higher error rates (Egner & Hirsch, 2005; Zinchenko et al., 2015). This may reflect increased competition for limited attentional resources (Etkin, Egner, Peraza, Kandel, & Hirsch, 2006). To some extent, the Virtual Apartment distractors appear to reflect these findings. However, future studies are needed to further differentiate the extent to which distractors in the Virtual Apartment are emotionally and/or socially salient. Moreover, as previously mentioned, there is need for future studies that offer a straightforward comparison between low and high distraction conditions within the virtual environment. A potential limitation of the current experimental design is that it relied on the computer automated Stroop that was completely devoid of environmental distractors for its nondistractor condition.

Another development that would improve the Virtual Apartment environment would be the addition of more function-led tasks that focus upon participant involvement in everyday activities (Parsons, 2015; Parsons, Carlew, Magtoto, & Stonecipher, 2017). The Virtual Apartment should be extended beyond traditional measures like the Stroop task that emphasize hypothetical constructs that may have little relevance to real-world behaviors. The next iteration of the Virtual Apartment could involve more "function-led" neuropsychological assessments that are developed from directly observable everyday behaviors (Burgess et al., 2006). Although virtual environments have been presented as potential aides in enhancing ecological validity, virtual environments like the Virtual Apartment Stroop was modelled on a construct-driven approach found in traditional Stroop assessments. Future iterations of the Virtual Apartment should include function-led tasks that represent the sorts of tasks needed for enhanced ecological validity and prediction of real-world functioning.

Furthermore, recent research indicates that healthy, undergraduate research subjects in a "neuropsychological experiment" may put forth suboptimal effort (An, Zakzanis, & Joordens, 2012). For the Virtual Apartment Stroop, we examine performance using scoring algorithms to screen for outliers and assess data integrity (Wu et al., 2013). It is important to note that the data are not free of potential confounds due to poor effort. Screening was done at the time of testing and during the data analysis process to eliminate obvious poor effort, but data may still contain some individuals who provided less than optimal effort. This situation is true for most normative neuropsychological data and particularly true for computerized testing given the reduced interaction with the examiner.

Future studies should make attempts at including well standardized effort tests with other cohorts (healthy

aging and age-related cognitive impairment) to assess the sensitivity and specificity of the Virtual Apartment Stroop Task. Furthermore, the findings are based on a relatively small sample size. As such, the results would be aided by a larger sample of participants. Although our findings need to be replicated to assess potential impact of sample size, effect size analysis revealed at least moderate effects for all analyses, which reduces the likelihood of anomalous results relative to sample size.

While the Virtual Apartment Stroop Task needs to be further validated, current findings provide data regarding the potential of the Virtual Apartment Stroop for fractionation of inhibition of prepotent responses and resistance to distractor inhibition, which is particularly impressive in our high-functioning sample. The Virtual Apartment Stroop replicated and extended the Stroop effects found in paper-and-pencil (e.g., D-KEFS) and computer-automated (e.g., ANAM) versions of the Stroop. Furthermore, the increased capacity of the Virtual Apartment Stroop Task to assess inhibitory control in the presence of real-world distractors may be useful in creating a greater amount of variability in performance levels of participants for investigations into inhibition of prepotent responses and resistance to distractor inhibition in real-world environments.

References

An, K. Y., Zakzanis, K. K., & Joordens, S. (2012). Conducting research with non-clinical healthy undergraduates: Does effort play a role in neuropsychological test performance? *Archives of Clinical Neuropsychology*, *27*, 849–857. doi:10.1093/arclin/acs085

Andrés, P., Guerrini, C., Phillips, L. H., & Perfect, T. J. (2008). Differential effects of aging on executive and automatic inhibition. *Developmental Neuropsychology*, *33*(2), 101–123. doi:10.1080/87565640701884212

Andres, P., & Van der Linden, M. (2000). Age-related differences in supervisory attentional system functions. *Journals of Gerontology, Series B. Psychological Sciences and Social Sciences*, *55*(6), 373–380.

Aretouli, E., & Brandt, J. (2010). Everyday functioning in mild cognitive impairment and its relationship with executive cognition. *International Journal of Geriatric Psychiatry*, *25*(3), 224–233. doi:10.1002/gps.2325

Armstrong, C. M., Reger, G. M., Edwards, J., Rizzo, A. A., Courtney, C. G., & Parsons, T. D. (2013). Validity of the Virtual Reality Stroop Task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology*, *35*, 113–123. doi:10.1080/13803395.2012.740002

Aschenbrenner, A. J., & Balota, D. A. (2015). Interactive effects of working memory and trial history on Stroop interference in cognitively healthy aging. *Psychology and Aging*, *30*, 1–8. doi:10.1037/pag0000012

Aschenbrenner, A. J., & Balota, D. A. (2017). Dynamic adjustments of attentional control in healthy aging. *Psychology and Aging*, *32*(1), 1–15. doi:10.1037/pag0000148

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, *49*, 5–28. doi:10.1080/027249896392784

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). San Diego, CA: Academic Press.

Bélanger, S., Belleville, S., & Gauthier, S. (2010). Inhibition impairments in Alzheimer's disease, mild cognitive impairment and healthy aging: Effect of congruency proportion in a Stroop task. *Neuropsychologia*, *48*(2), 581–590. doi:10.1016/j.neuropsychologia.2009.10.021

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 and Cognition*, *62*(2), 113–119. doi:10.1016/j.bandc.2006.04.006

Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age differences in Stroop interference: Contributions of general slowing and task-specific deficits. *Neuropsychology, Development, and Cognition*, *14*, 155–167. doi:10.1080/138255891007065

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L., Dawson, D. R., … Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*, 194–209.

Center for the study of human operator performance; Automated Neuropsychological Assessment Metrics (ANAM™): User manual. Norman, OK: C-SHOP, 2007.

Chafetz, M. D., & Matthews, L. H. (2004). A new interference score for the Stroop Test. *Archives of Clinical Neuropsychology*, *19*(4), 555–567. doi:10.1016/s0887-6177(03)00121-5

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201–216. doi:10.1016/j.acn.2007.08.010

Clapp, W. C., Rubens, M. T., Sabharwal, J., & Gazzaley, A. (2011). Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proceedings of the National Academy of Sciences*, *108*(17), 7212–7217. doi:10.1073/pnas.1015297108

CogState. (1999). CogSport [Computer software]. Parkville, Victoria, Australia: CogState, Ltd.

Collette, F., Schmidt, C., Scherrer, C., Adam, S., & Salmon, E. (2007). Specificity of inhibitory deficits in normal aging and Alzheimer's disease. *Neurobiology of Aging*, *30*(6), 875–889. doi:10.1016/j.neurobiolaging.2007.09.007

Coxon, J. P., Goble, D. J., Leunissen, I., Van Impe, A., Wenderoth, N., & Swinnen, S. P. (2014). Functional brain activation associated with inhibitory control deficits in older adults. *Cerebral cortex*, *26*(1), 12–22. doi:10.1093/cercor/bhu165

Davidson, D. J., Zacks, R. T., & Williams, C. (2003). Stroop interference, practice, and aging. *Aging, Neuropsychology and Cognition*, *10*, 85–98. doi:10.1076/anec.10.2.85.14463

Delis, D. C. (1997). Delis-Kaplan Executive Function Scale: Standardization edition manual. New York: The Psychological Corporation.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. Psychological Corporation.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784–1790. doi:10.1038/nn1594

Erlanger, D. M., Feldman, D. J., & Kutner, K. (1999). *Concussion resolution index*. New York, NY: HeadMinder.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*, 871–882.

Getzmann, S., Gajewski, P. D., & Falkenstein, M. (2013). Does age increase auditory distraction? Electrophysiological correlates of high and low performance in seniors. *Neurobiology of Aging*, *34*(8), 1952–1962. doi:10.1016/j.neurobiolaging.2013.02.014

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York, NY: Oxford University Press.

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. doi:10.1016/s0079-6123(07)00022-2

Henry, M., Joyal, C. C., & Nolin, P. (2012). Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: The bimodal virtual-reality Stroop. *Journal of Neuroscience Methods*, *210*, 125–131. doi:10.1016/j.jneumeth.2012.07.025

Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality of task set or degradation of task set across time? *Psychology and Aging*, *28*, 744–753. doi:10.1037/a0033159

Johnson, D. R., Vincent, A. S., Johnson, A. E., Gilliland, K., & Schlegel, R. E. (2008). Reliability and construct validity of the Automated Neuropsychological Assessment Metrics (ANAM) mood scale. *Archives of Clinical Neuropsychology*, *23*, 73–85. doi:10.1016/j.acn.2007.10.001

Kanske, P. (2012). On the influence of emotion on conflict processing. *Frontiers in Integrative Neuroscience*, *6*, 1–4.

Kim, S., Rasher, L., & Zacks, R. T. (2007). Aging and a benefit of distractibility. *Psychonomic Bulletin & Review*, *14*, 301–305. doi:10.3758/bf03194068

Lalonde, G., Henry, M., Drouin-Germain, A., Nolin, P., & Beauchamp, M. H. (2013). Assessment of executive function in adolescence: A comparison of traditional and virtual reality tools. *Journal of Neuroscience Methods*, *219*(1), 76–82. doi:10.1016/j.jneumeth.2013.07.005

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. doi:10.1016/j.tics.2004.12.004

Lovell, M. R., Collins, M. W., Podell, K., Powell, J., & Maroon, J. (2000). *ImPACT: Immediate post-concussion assessment and cognitive testing*. Pittsburgh, PA: NeuroHealth Systems.

Lustig, C., Tonev, S., & Hasher, L. (2000, April). Visual distraction and processing speed. In Poster presented at the Cognitive Aging conference, Atlanta, Georgia.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York, NY: Plenum.

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective, and social neurosciences. *Frontiers in Human Neuroscience*, 1–19.

Parsons, T. D., & Carlew, A. R. (2016). Bimodal Virtual Reality Stroop for assessing distractor inhibition in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *46*(4), 1255–1267.

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation*, *37*(5), 777–807. doi:10.1080/09602011.2015.1109524

Parsons, T. D., Courtney, C., Arizmendi, B., & Dawson, M. (2011). Virtual Reality Stroop Task for neurocognitive assessment. *Studies in Health Technology and Informatics*, *143*, 433–439.

Parsons, T. D., Courtney, C. G., & Dawson, M. E. (2013). Virtual Reality Stroop Task for assessment of supervisory attentional processing. *Journal of Clinical and Experimental Neuropsychology*, *35*, 812–826. doi:10.1080/13803395.2013.824556

Puccioni, O., & Vallesi, A. (2012). Conflict resolution and adaptation in normal aging: The role of verbal intelligence and cognitive reserve. *Psychology and Aging*, *27*, 1018–1026. doi:10.1037/a0029106

Reeves, D. L., Winter, K. P., Bleiberg, J., & Lane, R. L. (2007). ANAM Genogram: Historical perspectives, description, and current endeavors. *Archives of Clinical Neuropsychology*, *22*, S15–S37.

Rom, D. M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, *77*, 663–665.

Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology*, *8*, 81–88.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Thorne, D. R. (2006). Throughput: A simple performance index with desirable characteristics. *Behavior Research Methods*, *38*, 569–573. doi:10.3758/bf03193886

Uttl, B., & Graf, P. (1997). Color–Word Stroop Test performance across the adult life span. *Journal of Clinical and Experimental Neuropsychology*, *19*, 405–420. doi:10.1080/01688639708403869

Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. Psychological Corporation. San Antonio, TX.

Wu, D., Lance, B., & Parsons, T. D. (2013). Collaborative filtering for brain-computer interaction using transfer learning and active class selection. *PLoS ONE*, *8*(2), 1–18.

Xue, S., Cui, J., Wang, K., Zhang, S., Qiu, J., & Luo, Y. (2013). Positive emotion modulates cognitive control: An event-related potentials study. *Scandinavian Journal of Psychology*, *54*(2), 82–88. doi:10.1111/sjop.12031

Zinchenko, A., Kanske, P., Obermeier, C., Schröger, E., & Kotz, S. A. (2015). Emotion and goal-directed behavior: ERP evidence on cognitive and emotional conflict. *Social Cognitive and Affective Neuroscience*, *10*(11), 1577–1587.