Sustained attention across the lifespan in a sample of 10,000: Dissociating ability and strategy

Francesca C. Fortenbaugh[a,b,*], **Joseph DeGutis**[a,c], **Laura Germine**[d,e], **Jeremy Wilmer**[f], **Mallory Grosso**[a], **Kathryn Russo**[a], and **Michael Esterman**[a,g]

[a] Department of Veterans Affairs, Boston Healthcare System, Boston, MA 02130

[b] Department of Psychiatry, Harvard Medical School, Boston, MA 02115

[c] Department of Medicine, Harvard Medical School, Boston, MA 02115

[d] Department of Psychology, Harvard University, Cambridge, MA 02138

[e] Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114

[f] Department of Psychology, Wellesley College, Wellesley, MA 02481

[g] Department of Psychiatry, Boston University School of Medicine, Boston, MA 02118

Abstract

Normal and abnormal differences in sustained visual attention have long been of interest to scientists, educators, and clinicians. Still lacking, however, is a clear understanding of how sustained visual attention varies across the broad sweep of the human lifespan. Here, we fill this gap in two ways. First, powered by an unprecedentedly large, 10,430-person sample, we model age-related differences with substantially greater precision than prior efforts. Second, using the recently developed gradual-onset continuous performance test (gradCPT), we parse sustained attention performance over the lifespan into its ability and strategy components. We find that after age 15, the strategy and ability trajectories saliently diverge. Strategy becomes monotonically more conservative with age, whereas ability peaks in the early forties and is followed by a gradual decline in older adults. These observed lifespan trajectories for sustained attention are distinct from results of other lifespan studies focusing on fluid and crystallized intelligence.

Keywords

Sustained attention; Response Inhibition; Development; Aging

Introduction

The ability to sustain attention to a task over seconds to minutes is a core cognitive function that plays a critical role in daily functioning. For example, sustained attention has been linked to academic and employment performance (Kalechstein, Newton, & van Gorp, 2003; Lam & Beale, 1991) and attentional lapses predict driving accidents (Ball, Owsley, & Sloane, 1991; Edkins & Pollock, 1997; Schmidt, et al., 2009). Research has also shown that numerous other cognitive processes rely on sustained attention, such as learning, memory and executive functions (Barkley, 1997; Sarter, Givens, & Bruno, 2001; Silver & Feldman, 2005). Furthermore, deficits in sustained attention are one of the most pervasive cognitive issues across a wide range of neurological and psychiatric patient populations (Buxbaum, et al., 2004; Clark, Iversen, & Goodwin, 2002; Park, Hood, Shah, Fogg, & Wyatt, 2012). Despite numerous investigations of how other cognitive processes change across the lifespan (e.g., verbal ability or processing speed), lifespan changes in sustained attention remain to be fully characterized. Filling this gap will not only inform cognitive models of human development and aging, but will also help better define and pinpoint the mechanisms of cognitive dysfunction in neurologic and psychiatric populations.

The studies completed thus far on age-related changes in sustained attention ability have yielded inconsistent results (Staub, Doignon-Camus, Després, & Bonnefond, 2013), ranging from poorer performance in both childhood and aging relative to adulthood (McAvinue, et al., 2012), to no changes between younger and older adults (Bunce & Sisa, 2002; Staub, et al., 2013), to improved functioning in older adults (Carriere, Cheyne, Solman, & Smilek, 2010; Staub, Doignon-Camus, Bacon, & Bonnefond, 2014). One potential cause for the disparate findings is that by focusing on error rates and not comparing hits vs. false alarms, previous studies have not dissociated changes in strategy and ability across the lifespan (Sarter, et al., 2001). Moreover, the primary type of errors made can vary with the task (commissions versus omissions), making comparisons difficult (Staub, et al., 2013).

Recently, our laboratory has developed the gradual-onset continual performance task (gradCPT), with the aim of better characterizing individual differences in sustained attention (Esterman, Noonan, Rosenberg, & DeGutis, 2012; Rosenberg, Noonan, DeGutis, & Esterman, 2013). The gradCPT represents a unique combination of task features, in that it both requires frequent overt responses and removes abrupt stimulus onsets that exogenously capture attention. Requiring frequent overt responses is common in continuous performance tasks (e.g., Sustained Attention to Response Task, Robertson, Manly, Andrade, Baddeley, & Yiend, 1997) and allows for reliable analyses of response timing and variability, and accuracy across the whole task as well as within periods of high and low attentional stability (in versus out of the zone). Additionally, the use of gradual stimulus changes makes performance less tied to phasic stimulus onsets and offsets, better isolating intrinsic sustained attention abilities. To separate the contribution of strategic changes and ability factors, we have also successfully utilized signal detection analyses of gradCPT (Esterman, Reagan, Liu, Turner, & DeGutis, 2014). Finally, several recent studies suggest that the gradCPT is ecologically valid in that a variant of the gradCPT correlates with real-world attentional problems (Rosenberg, et al., 2013) and performance on the original version is impaired in patient populations with traditionally poor sustained attention (Auerbach, et al.,

2014; DeGutis, et al., in press). Together, this suggests the gradCPT is a powerful tool to capture changes in sustained attention ability throughout the lifespan.

To help resolve how sustained attention changes across the lifespan, the current study tested an unprecedented sample of 10,430 participants between 10-70 years old on an adapted 4-minute version of the gradCPT (Esterman, et al., 2012). This sample is larger than all previous efforts to model changes in sustained attention performance during development, aging, or across the lifespan combined, allowing us to more precisely model transition periods in performance across the lifespan using segmented linear regression. We also employed factor analyses and confirm the existence of two latent, dissociable factors underlying gradCPT performance – the ability to sustain attention (discrimination performance and response time consistency) and the strategic approach (response speed and carefulness). The results show unique patterns in how the ability and strategy factors change across the lifespan, and suggest that the lifespan trajectory of sustained attention ability is unique from the trajectories other studies have found for crystallized intelligence (e.g., vocabulary), which continues to improve throughout the lifespan until the mid 60s, as well as fluid intelligence (e.g., working memory), which peaks in the mid-twenties (Craik & Bialystok, 2006; Hartshorne & Germine, in press).

Methods

Participants

Ten thousand four hundred thirty unpaid volunteers, aged 10 to 70 (M=26.07, SD=11.77), were included in the analyses of this online study (Figure 1). These participants were visitors from March 2014 to September 2014 of TestMyBrain.org, a cognitive testing website that provides feedback on performance relative to other users. The gradual onset continuous performance task (gradCPT) was presented on the home page of TestMyBrain.org as a "Continuous Concentration task". TestMyBrain.org receives traffic mostly from social networking sites and search engines (Germine, et al., 2012). Participants were asked at the end of the experiment: "Was this your first time participating in this particular research experiment?" Data from repeat participation was excluded. Within seven months, 10,922 people completed the task. Of these participants, 342 were excluded due to missing data or technical problems. Amongst technical problems, we chose to exclude those computers that exceeded 10% error in the average stimulus presentation time, meaning the time it took for a new image to transition from 0% to 100% opacity. This included average stimulus presentation times that were too fast (720ms) or too slow (880ms). From the remaining 10,580 participants another 150 (1.4%) were excluded for "tune-outs", defined as a 30sec interval or more without a response. Of the 10,430 participants that were included, there was a nearly equal ratio of males and females (5,027 males and 5,403 females).

Task and Procedure

The gradCPT is one test out of several on TestMyBrain.org (e.g., face recognition, working memory). Participants are free to complete one or more of these tasks meaning that for some participants the gradCPT was the first or only test participants completed, while others may have completed other tasks on TestMyBrain prior to completing the gradCPT. Single

experiment studies on the TestMyBrain.org website are kept brief (< 10 minutes) in order to maintain a balance between task completion, participation, and the test length. Given the demanding nature of continuous performance tasks, the web-based gradCPT used a shortened (4 min vs. 8 min) version of the continuous go-no go task originally reported in Esterman et al. (2012). This test length was chosen because it was sufficiently short such that participant attrition rates were comparable to other experiments on the site. The concern with participants dropping out during testing regards a selection bias in which individuals with poorer sustained attention ability would be less likely to complete the experiment. This led to a total experiment time of approximately 7 minutes from consent to debriefing.

The stimuli consisted of 20 round (200 pixel diameter), grayscale photographs of 10 city scenes and 10 mountain scenes. The same trial sequence was used for every participant. This approach is regularly used in individual differences research (Carlson & Moses, 2001) in order to eliminate the order of stimuli as a potential source of systematic variation across age groups. The order of scenes was pseudo-random (90% cities and 10% mountains) in that identical scenes did not appear on consecutive trials. Each scene gradually transitioned to the next scene every 800 ms using a linear pixel-by-pixel interpolation, for a total of 299 trials across the 4 minutes of testing. Additionally, the block began with a fully opaque scrambled city image from which the first trial image transitioned into and the same scrambled image was used at the very end of the block for the last trial image to transition into. Figure 2 illustrates the linear interpolation utilized. At the start of every new trial, the incoming scene has an image opacity of 0% (i.e. is 100% transparent) and transitions to 100% opacity by the end of the 800 ms. On the following trial this scene then transitions from 100% opacity back to 0% opacity while a new image transitions into view. Participants were instructed to press the spacebar for city scenes (267 trials) and withhold a response for mountain scenes (32 trials). Thus, the task instructions emphasized accuracy in responses while the continuous nature of the task imposed a limited time within which participants could respond.

Before starting the gradCPT, participants gave informed consent according to the guidelines set by the Institutional Review Board Committees for the Use of Human Subjects at Harvard University and Wellesley College. Participants were then asked to complete a voluntary demographic survey that included age, gender, native language, and ethnicity. Comprehensive instructions as well as 3 practice sessions (30 seconds each) were then provided to familiarize the participants with the scenes and the task so that it could be completed without assistance from an experimenter. After completing 4-mintues of the gradCPT, participants were asked if they had cheated or if any problems occurred during the task and then they were provided with performance feedback. This feedback includes a personal score, which is the percentage of correct withholds to mountains, as well as how this score compares to the average participant.

Analyses

An iterative algorithm was used to assign button presses to each trial, following the methodology outlined in Esterman et al. (2012). Reaction times were defined relative to the beginning of each image transition, such that a reaction time of 800 ms indicates a response

when the current trial image reached 100% opacity while shorter reaction times indicate that the current trial image was still transitioning when a response was made. All responses were logged throughout the experiment and the algorithm sets a limit of 1,360ms before a response time is assigned to the next trial (see Supplementary Methods available online for more details).

Using correct responses to cities, we computed the mean reaction time (ms) and reaction time variability. Reaction time variability was calculated using the coefficient of variation (CV), or the standard deviation of reaction times divided by the mean reaction time for each participant. Commission error rate (pressing to a target/mountain scene) and omission error rate (failing to press to a non-target/city scene) were then calculated.

Additionally, taking a signal detection approach with the hits (correct omissions to mountains) and false alarms (incorrect omissions to cities), we computed d' (a measure of discrimination ability) and criterion (a measure of strategy/willingness to respond in the case of uncertainty). We used standard procedures to correct for cases when hit rates were 100% or false alarm rates were 0% with one-half error deducted or added based on the number of target or non-target trials presented, respectively (see Supplementary Methods available online). Factor analyses were completed with a direct oblimin rotation. This uses an oblique rotation that is more accurate than an orthogonal rotation, and provides a more optimal solution when the resulting factors are correlated. Commission and omission errors were not used in the factor analysis due to non-independence with the signal detection variables.

An additional feature of the gradCPT is the analysis procedure of splitting the duration of the task into states of low and high variability (in the zone vs. out of the zone). Although this version was abbreviated to 4 min (and thus limits in vs. out of the zone epochs to 2-min each), we nonetheless explored performance during each of these states using the variance time course (VTC) analysis (Esterman, et al., 2012). Specifically, the reaction times for correct responses were converted to absolute z-scores and values for correct omissions and error trials were linearly interpolated using the average of the two neighboring trials. The VTC was then smoothed using a Gaussian kernel with a 20 trial window and an 8 second full width at half maximum. From the smoothed VTC for each participant, two attentional states were defined using a median split to separate epochs of low or high variability ("in the zone" and "out of the zone" epochs, respectively). Thus, in-the-zone epochs include trials where reaction times were closest to the mean of the run while out-of-the-zone epochs include trials with the most deviant reaction times including both the fastest and the slowest responses.

Results

The following analyses focus on four dependent measures of interest: reaction time, reaction time variability, d', and criterion (see Supplementary Figure S3 available online for commission and omission error results).

Reliability of dependent measures

Reliability measurements were obtained using Spearman-Brown corrected split-half correlations comparing the averaged performance of the 1st and 3rd minutes with the averaged performance of the 2nd and 4th minutes. All four dependent measures show acceptable to high internal reliability: mean reaction time ($rel_{sb} = 0.94$), reaction time variability ($rel_{sb} = 0.90$), d' ($rel_{sb} = 0.78$), and criterion measures ($rel_{sb} = 0.80$).

Comparison of lab-based and web-based measures

To test whether the gradCPT performed similarly over the web as it does in the lab, we selected the subset of 6,290 participants between the ages of 18-34 years and compared their performance to the first four minutes of gradCPT data in 17 age-matched participants previously collected in a controlled laboratory setting (Esterman, et al., 2012) (see Supplementary Figure S1 available online). Independent t-tests assuming unequal variance show no difference in accuracy (d'), criterion, commission or omission error rates (p 0.29 for all). Mean reaction time is slightly slower in the present sample ($X_{diff} = 81ms$; $t(6305) = 5.43$, $p<0.001$, d = 1.20) which may be due to technical/hardware differences in response collection over the Internet versus the laboratory (McGraw, Tew, & Williams, 2000). Additionally, the reaction time variability was lower in the web-based sample ($X_{diff} = -0.029$; $t(6305) = 2.44$, $p=0.01$, d = 0.56). Importantly, however, the differences in reaction time measures do not accompany differences in accuracy (d', commission or omission error rates) or criterion. This supports alignment between the performance of young adults in controlled laboratory settings with the web-based participants in the present study and is consistent with previous web-based studies showing comparable performance levels to lab-based studies of perception, attention, and working memory (Germine, et al., 2012; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Hartshorne & Germine, in press; McGraw, et al., 2000).

Modeling lifespan changes in sustained attention performance

To achieve our goal of understanding the mechanisms of sustained attention changes across the lifespan, we separately modeled each of our four key sustained attention measures across the 10-70 year age range (see Figure 3). At a descriptive level, one could expect three types of processes across time: a growth in capability, a plateau where capability is maintained, and declines in capability. To uncover these possible changes, we modeled the data using the average data for each age. As the age distribution of the participants is skewed (Figure 1), we further binned the youngest and oldest participants into age bins with the constraint that at least 100 participants fall into each bin, with the mean age used in the following analyses. The binned age groups were: 10-12, 33-34, 35-36, 37-38, 39-40, 41-43, 44-46, 47-49, 50-53, 54-57, 58-61, and 62-70 years. We note that very similar results are found when the age means are modeled without binning and with the raw individual participant data (see Supplementary Results available online). We first applied a standard polynomial approach previously used to model lifespan changes in sustained attention (McAvinue, et al., 2012). Polynomial modeling revealed, for all four measures, significant non-linearity (cubic trends p 0.0003 for all).

To quantify the nature and timing of performance changes across the lifespan, we next employed hierarchical regression analyses utilizing segmented linear functions (i.e. "piecewise regression"). The benefit of these segmented linear functions is the ability to capture multiple combinations of changes (growth, plateau, decay) with no assumption of symmetry such as that implicitly assumed when utilizing quadratic functions in trend analyses. Moreover, the breakpoint between the two linear segments and the confidence intervals around these breakpoints provide a direct estimate of a "transition zone," or an age range where transitions are most likely to occur. Segmented linear functions have been successfully used to model lifespan changes in white matter tract integrity (Yeatman, Wandell, & Mezer, 2014). However, to date, such models have not been applied to lifespan changes in cognition. In the present study, linear functions were compared to a two-phase segmented linear function with one transition (breakpoint). If the two-phase model provided a significantly better fit, we then compared two-phase functions with three-phase functions (1 vs. 2 transitions).

Changes in ability across the lifespan

As seen in Figure 3, d' and reaction time variability measures exhibited a similar pattern across the lifespan with both being fit best by 3-phase linear functions (see Table 1 for model comparisons), showing rapid development in sustained attention ability between 10-16 years of age, then a period of relative stability until ~43 years of age, and finally a decline in ability across old age. In particular, the estimated breakpoints for d' occur at 16.5 years (95% CI: 15.9 to 17. 1) and 42.9 years (95% CI: 40.3 to 45.5), which are notably similar to the reaction time variability breakpoints at 16.4 years (95% CI: 15.9 to 16.9) and 44.3 years (95% CI: 41.2 to 47.4).

Examining the slopes, or rate of change during each phase, discrimination ability (d') rapidly improves between 10-16 years (0.13/year; 95%CI = 0.11 to 0.16). This is followed by a period of relative stability with modest but significant increases in discrimination ability between 17-43 years of age (0.012/ year; 95%CI = 0.0097 to 0.015; $t(26) = 9.71$, $p < 0.0001$). D' peaks at ~43 years and following this, a gradual decline is observed (−0.018/yr; 95%CI = −0.023 to −0.013). Similarly, for reaction time variability, the slope patterns highlight a rapid performance improvement (decrease in response variability) between 10-16 years of age (decrease in coefficient of variation by −0.01ms/ year; 95%CI: −0.011 to −0.008). This is followed by a period of stability between 16-44 years of age, with small but unreliable decreases in reaction time variability (−0.0001 per year; 95% CI: −0.0003 to 0.00006; $t(26) = 1.35$, $p = 0.19$). Performance then declines beyond age 44 with reaction time variability increasing at a rate of 0.0014/year (95%CI: 0.001 to 0.002). For both measures, the estimated slopes during older adulthood are less than 1/5 those observed during the childhood development period, suggesting that while participants show a decline in task ability as they get older, the rate of decline observed is not nearly as great as the rate of increase in task ability seen during development. Importantly, beyond the rapid development observed between 10-16 year olds, the results show a continued though slowed increase in discrimination ability (d') with maximum sustained attention ability occurring far later in life in the mid 40's. This result highlights a unique lifespan trajectory for sustained attention ability compared to other cognitive abilities, such as fluid intelligence, which

studies have shown peaks in the mid 20's (Craik & Bialystok, 2006; Hartshorne & Germine, in press).

Changes in strategy across the lifespan

The criterion and mean reaction time measures show markedly distinct lifespan patterns from d' and reaction time variability (see Figure 3), though these measures were also best fit by the 3-phase linear functions (see Table 1 for model comparisons). In particular, the mean reaction time measures and criterion measures show a period of rapid development followed by two distinct but monotonically changing shifts in strategy across the latter years. For the mean reaction time measure, the estimated breakpoints occur at 15.0 years (95% CI: 14.0 to 16.6) and 58.5 years (95% CI: 55.1 to 61.9). The slope patterns highlight a trend toward a speeding of reaction time between 10-15 years of age (speeding at −3.45ms/ year; 95%CI: −7.1 to 0.1; $t(26) = 1.97$, $p = 0.06$). Following this period a reversal is observed with reaction times slowing between 15-58 years of age at a rate of 2.7ms per year (95% CI: 2.6 to 2.9). After the second transition point, reaction times flatten with the slope estimate showing an unreliable speeding rate of −0.6ms per year (95% CI: −2.5 to 1.3; $t(26) = 0.60$, $p = 0.52$). The slowing of reaction times from 16-29 years of age likely reflects a strategic shift, since choice reaction time has been shown to speed up in this range (Williams, Hultsch, Strauss, Hunter, & Tannock, 2005). The monotonic increase in reaction times observed from 30-58 years of age could also reflect a strategic shift, though it is also consistent with a general age-related slowing of reaction times (Ratcliff, Thapar, & McKoon, 2001). The change in slope observed beyond the second breakpoint may be related to the experiment design and the implicit reaction time limitations inherent in continuous performance tasks (i.e., reaction time ceiling effect). Specifically, the algorithm used to assign reaction times has an implicit maximum of 1,360ms before responses are assigned to the next trial. While no individual participant showed mean reaction times close to or above this ceiling (see Supplementary Methods available online), it is possible that more reaction times above this maximum occurred in participants over 58 years of age leading to an increase in reaction time variability and a flattening of the calculated mean reaction times.

The pattern of change observed for the criterion measure shows similarities to that observed for the mean reaction time measure. First, there is a slight upward slope toward a greater bias to respond (more impulsive strategy) in the first phase with an initial breakpoint at 14.5 years (95% CI: 13.0 to 16.9), though this slope is not statistically greater than zero (slope = 0.003; 95% CI: −0.01 to 0.02; $t(26)=0.43$, $p = 0.67$). Following the initial transition point, indicating the least cautious approach to the task, a change in the pattern occurs and similar to the reaction time measure a monotonic trend is observed across the rest of the age groups suggesting a continuous shift toward a more cautious approach to the task reflected in a reduced bias to press on a given trial. In contrast to the other performance measures, the second transition point is seen earlier at 27.1 years (95% CI: 24.4 to 29.8). Between 15-27 years a reliable shift toward a more conservative approach is observed with a slope of −0.01/ year (95% CI: −0.020 to −0.01). Following the second transition zone, the rate of change slows but continues toward a decreased bias to press to a given trial (slope = −0.0052/year; 95%CI: −0.006 to −0.004).

Collectively, these measures show that beyond ~15 years of age, participants show a gradual slowing of reaction times and a more careful approach to the task with age. This is in line with previous studies demonstrating more impulsivity in adolescence (Galvan, et al., 2006) and the observed slowing of reaction time from 15-30 years old when reaction time abilities are typically still improving (Williams, et al., 2005) is suggestive of strategic slowing. In contrast, the slowing of reaction times from ages 30-58 likely reflects a combination of decreasing simple target detection reaction time (Ratcliff, et al., 2001) and a more cautious strategy, which has previously been demonstrated in older adults (Deakin, Aitken, Robbins, & Sahakian, 2004). To further investigate and provide converging evidence for this strategic shift across the lifespan, we calculated the degree to which individuals slowed down responses following an incorrect response. This phenomenon, referred to as post-error slowing, is thought to reflect error monitoring (Dutilh, et al., 2012). To better isolate post-error slowing, across the group we regressed out the mean reaction time for correct trials immediately preceding an error trial from the mean reaction times for correct trials immediately following an error trial. As seen in Figure 4, post-error slowing shows a very similar lifespan pattern as criterion and reaction time, consistent with the interpretation that the three measures are driven by strategic shifts towards cautiousness. The pattern of post-error slowing data was best fit by a 2-phase segmented function (see Table 1). Similar to the criterion and reaction time results, the parameters of the 2-phase model show a transition zone at 15.0 years (95%CI: 13.3 to 16.8), representing the lowest error monitoring in early adolescence. As with the criterion measure, in the first phase a negative but unreliable slope is observed (slope = −4.36; 95%CI: −9.9 to 1.1; $t(28)=-1.62$, $p = 0.12$). However, following the transition zone the results show a consistent increase in post error slowing of 3.5ms per year (95%CI: 3.2 to 3.7). This result suggests a gradual shift towards greater error monitoring across the lifespan above and beyond that explained by slower reaction times due to changes in central nervous functioning in older adults (Ratcliff, et al., 2001), and is consistent with previous self-reports by older participants of less mind wandering during task completion and greater intrinsic motivation to perform well on similar tasks (Staub, et al., 2014). Collectively, then, the mean reaction time, criterion, and post-error slowing measurements all show a monotonic trend across participants 15 years and older representing a strategic shift toward a slower, more cautious approach to the task that diverges significantly from the pattern observed in the d' and reaction time variability measures.

Factor analysis of latent measures in sustained attention performance

To further confirm the dissociation between sustained attention ability and strategy, we conducted exploratory factor analyses on the four primary variables (mean reaction time, reaction time variability, d', criterion). Two components emerged that have eigenvalues over Kaiser's criterion of 1: 1) reaction time variability and d' and 2) mean reaction time and criterion (see Figure 5a and Table 2). Together these components explain 85.21% of the variance. Given the transition zones detected in sustained attention ability across the lifespan, we also conducted the factor analysis for each age group based on the break points observed for the RT variability and d' factors (10-16, 17-43, 44-70). The same variables cluster together across all three age ranges providing further support for two latent variables relating strategy and ability to task performance and changes in task performance across the

lifespan. Finally, re-running the factor analysis when including post-error slowing shows very similar results, with post-error slowing clustering with reaction time and criterion while d' and reaction time variability load onto a separate component (Figure 5b). Together, these results further confirm that d' and reaction time variability measure a similar latent variable (i.e., sustained attention ability) which is dissociable from the latent variable that criterion, reaction time, and post-error slowing measure (i.e., sustained attention strategy).

Comparison of lifespan changes in and out of "the zone"

In our final analysis we explored whether the observed age-related changes in the task as a whole are consistent across states of high and low attentional stability (in- and out-of-the-zone, see Methods). During in-the-zone (low variability) and out-of-the-zone (high variability) epochs, we compared two variants of the 3-phase segmented linear functions used in the previous analyses. In this analysis, we model a simple main effect of zone as a shift in the intercept parameter (i.e. baseline shift) across the two zone conditions while an interaction is reflected in a shift in any of the other five parameters (2 breakpoints, 3 slopes). Thus, the null model was a model in which the intercept parameters were free to vary across the two zone conditions while the other five parameters were shared. In the alternative model all 6 parameters were free to vary across the two zone conditions. Qualitatively, the developmental trajectories are similar across attentional states with overlapping transition points and the same slope directions when participants were in the zone and out of the zone. However, results of the model comparison show that only the criterion measure was not significantly better fit by the alternative model, $F(5,52) = 1.42$, $p = 0.23$, indicating that the change in criterion when participants are in-the-zone versus out-of-the zone is well modeled by an additive shift of -0.095 for all ages (see Figure 6 and Table 3). In contrast, all three of the remaining performance variables were better fit by the alternative model indicating some differences in lifespan trends across the zone conditions (mean reaction time: $F(5,52) = 3.83$, $p = 0.005$; reaction time variability: $F(5,52) = 27.50$, $p < 0.0001$; d': $F(5,52) = 3.32$, $p = 0.01$). Inspection of Figure 6 and Table 3 shows that while these three performance variables show similar trajectories across the age ranges tested, the most prominent differences occur during the development and aging phases. In particular, the d' slope for participants over the age of 43 years increases by ~78% when they are out-of-the-zone compared to in-the-zone indicating that age related declines in task sensitivity are more prominent during times of attentional instability. Similarly, a two-fold increase is seen in the slope parameter for reaction time variability in participants over the age of 44 years indicating that age-related increases in reaction time variability are more prominent when participants are out-of-the-zone than in-the-zone. Additionally, the differences in the initial slope parameters for the reaction time and reaction time variability measures suggest that differential developmental changes may occur in response generation across periods of low and high attentional stability.

Discussion

The present study uses an unprecedented sample size and novel methods to explore sustained attention across a sixty-year lifespan. We demonstrate that two distinct underlying processes contribute to sustained attention performance: 1) the ability to maintain consistent

and accurate performance and 2) the strategy of going faster with a bias to respond, or slower with a bias to withhold.

The modeling results show that these two underlying factors have differential lifespan trajectories with critical transitional phases. All performance measures suggest an initial period of development with a rapid increase in task ability and shift toward a faster, less cautious strategy across the youngest age groups, with early transition zones in adolescence around 14-17 years of age. However, beyond this initial transition period important dissociations in lifespan trajectories are observed with task ability parameters (d' and reaction time variability) showing evidence for continued improvement through adulthood and marked decreases in ability beyond 43 years of age. In contrast, adult participants show a monotonic trend in strategy changes, toward a slower and more conservative approach to the task. Different strategy life phases are best characterized by changes in the magnitude of the slope parameters (i.e. the rate of change) but not in the direction. This shift in strategy was further supported by age-related changes in post-error slowing, which again highlight a monotonic shift toward increased error monitoring beyond age 15. While simple RTs are known to slow in older adults due to primary sensorimotor changes (Ratcliff, et al., 2001), the continuous nature of the changes observed in reaction time and criterion measures from as young as 16 years of age, coupled with the same pattern observed in the post-error slowing measure when controlling for overall reaction time, suggest that the changes across the full age range is best accounted for by changes in task strategy. This does not discount any role for sensorimotor changes in influencing these measures. One intriguing possibility is that the relative influence of such factors increases with age and ultimately limits the benefit of strategic slowing in the older participants with regards to task accuracy.

The dissociable lifespan trajectories between strategy and ability were also evident during both participants' relative best (in-the-zone) and worst (out-of-the-zone) periods of performance. The most notable difference between attentional states is the decline in the ability parameters after ~43 is markedly steeper during out-of-the-zone periods. Such periods are thought to reflect the most taxing periods of the task (Esterman, Rosenberg, & Noonan, 2014), and suggests more pronounced age-related decline in more challenging tasks. The results of factor analyses further demonstrated that the dissociation between strategy and ability is evident across the sample as a whole, as well as within each of the three specific age-ranges tested.

The use of segmented linear regression analyses points to important regions of time that we have labeled transition zones. While the class of functions we utilize here defines a specific breakpoint for each zone, changes are likely to occur gradually over a period of time in individuals. One particularly useful aspect of these functions is that they provide explicit estimates of transition periods in the parameter estimates which, when interpreted with confidence intervals, provide a likely time-window across which transition zones may be expected and developmentally appropriate. Such normative data could provide a basis for potentially revealing abnormal lifespan trajectories, such as those associated with developmental disorders (e.g., ADHD) and pathological aging (e.g., dementia).

More broadly, the results also suggest that sustained attention ability peaks far later in life than other visual and cognitive processing mechanisms. Specifically, the results show that despite a slowing in growth during adulthood, sustained attention ability (i.e., d') peaks at 43 years of age. This is far later than other cognitive abilities such as those related to fluid intelligence. Studies have shown that sensory and cognitive processing abilities, including visual processing speed and working memory, peak before age 30 and decline thereafter (Baltes & Lindenberger, 1997; Germine, Duchaine, & Nakayama, 2011; Halberda, et al., 2012; Hartshorne & Germine, in press; Owsley, 2011). Conversely, sustained attention ability shows an earlier peak than has been measured for simple knowledge accumulation related to crystallized intelligence (Craik & Bialystok, 2006; Hartshorne & Germine, in press). This suggests that sustained attention ability represents an important, distinct mechanism that contributes to an individual's ability to process information and interact with the world. While young adults may surpass others in the speed and flexibility of information processing, and older adults may possess the most stored knowledge regarding the world, we find that middle-aged adults have the greatest capacity to remain "attentive". One explanation for sustained attention ability peaking at 43 is that attention is highly trainable (DeGutis & Van Vleet, 2010), and practice focusing attention throughout adulthood may further hone this skill.

Sustained attention peaking in one's 40's is also consistent with recent studies of white matter and prefrontal cortex integrity across the lifespan (Hedden & Gabrieli, 2004; Yeatman, et al., 2014). One recent study has shown asymmetrical maturation and degeneration processes in frontal white matter tract integrity across the lifespan which qualitatively matches the pattern observed in our ability factor (Yeatman, et al., 2014). Sustained attention activates a large-scale network of cortical and subcortical regions, including areas in the frontal lobe (Esterman, et al., 2012). Thus, changes in frontal white matter tract integrity over time may significantly impact many cognitive functions, including sustained attention ability.

There are also several limitations of the present study that future studies may address. First, the present study used a cross-sectional design, preventing assessment of individual lifespan trajectories. Further research utilizing longitudinal studies would provide more useful data for investigating the pattern of changes within these transition zones. Second, participants were free to potentially complete multiple experiments before participating in the current experiment. This most likely introduced inter-participant variability rather than bias across the age ranges, but future studies may wish to account for this potential issue. Third, the present sample is skewed, with a relative under-sampling of the youngest and oldest participants. As the modeling results were consistent regardless of which data was used (binned/un-binned age means or individual data), this is unlikely to have greatly impacted the results as the smallest binned age group still had 122 participants. A related issue is the potential under-estimation of age-related declines. Given the use of internet-based volunteers, it is possible a selection bias occurred in the older participants towards higher-functioning older adults. While other TestMyBrain.org studies have replicated lifespan trends in cognition compared to traditionally-collected and U.S. nationally-representative samples (Hartshorne & Germine, in press), future research obtaining greater demographic

and educational details, along with extending the sample age range, would help to validate the current findings.

Despite these limitations, the current results provide important information that will help to fill in the gaps in our understanding of normative changes in sustained attention across the lifespan. Utilizing factor analyses and a novel regression approach that highlights important transition periods in sustained ability and the strategies utilized by participants these results provide a new foundation for future research on sustained attention as well as studies on a range of neurocognitive functions that depend on sustained attention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Auerbach RP, Kim JC, Chango JM, Spiro WJ, Cha C, Gold J, et al. Adolescent nonsuicidal self-injury: Examining the role of child abuse, comorbidity, and disinhibition. Psychiatry Research. 2014; 220(1-2):579–584. [PubMed: 25095754]

Ball K, Owsley C, Sloane ME. Visual and cognitive predictors of driving problems in older adults. Experimental Aging Research. 1991; 17:79–80. [PubMed: 1794385]

Baltes PB, Lindenberger U. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cogntiive aging? Psychology and Aging. 1997; 12(1):12–21. [PubMed: 9100264]

Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychological Bulletin. 1997; 121(1):65–94. [PubMed: 9000892]

Bunce D, Sisa L. Age differences in perceived workload across a short vigil. Ergonomics. 2002; 45(13):949–960. [PubMed: 12519526]

Buxbaum LJ, Ferraro MK, Varamonti T, Farne A, Whyte J, Ladavas E, et al. Hemispatial neglect: Subtypes, neuroanatomy, and disability. Neurology. 2004; 62:749–756. [PubMed: 15007125]

Carlson SM, Moses LJ. Individual differences in inhibitory control and children's theory of mind. Child development. 2001; 72(4):1032–1053. [PubMed: 11480933]

Carriere JSA, Cheyne JA, Solman G, Smilek D. Age trends for failures of sustained attention. Psychology and Aging. 2010; 25(3):569–574. J.F. [PubMed: 20677878]

Clark L, Iversen SD, Goodwin GM. Sustained attention deficit in bipolar disorder. The British Journal of Psychiatry. 2002; 180(4):313–319. [PubMed: 11925353]

Craik FIM, Bialystok E. Cognition through the lifespan: mechanisms of change. Trends in Cognitive Science. 2006; 10(3):131–138.

Deakin J, Aitken M, Robbins T, Sahakian B. Risk taking during decision-making in normal volunteers changes with age. Journal of the International Neuropsychological Society. 2004; 10:590–598. [PubMed: 15327737]

DeGutis JM, Esterman M, McCulloch B, Rosenblatt A, Milberg W, McGlinchey R. Post-traumatic psychological systems are associated with reduced inhibitory control, not general executive dysfunction. Journal of the International Neuropsychological Society. in press.

DeGutis JM, Van Vleet TM. Tonic and phasic alertness training: A novel behavioral therapy to improve spatial and non-spatial attention in patients with hemispatial neglect. Frontiers in Human Neuroscience. 2010; 4(60):117.

Dutilh G, Vandekerckhove J, Forstmann BU, Keuleers E, Brysbaert M, Wagenmakers E-J. Testing theories of post-error slowing. Attention, Perception, & Psychophysics. 2012; 74:454–465.

Edkins GD, Pollock CM. The influence of sustained attention on Railway accidents. Accident Analysis & Prevention. 1997; 29(4):533–539. [PubMed: 9248512]

Esterman M, Noonan SK, Rosenberg M, DeGutis J. In the Zone or Zoning out? Tracking behavioral and neural fluctuations during sustained attention. Cerebral Cortex. 2012; 23(11):2712–2723. [PubMed: 22941724]

Esterman M, Reagan A, Liu G, Turner C, DeGutis J. Reward reveals dissociable aspects of sustained attention. Journal of Experimental Psychology: General. 2014; 143(6):2287–2295. [PubMed: 25313950]

Esterman M, Rosenberg M, Noonan SK. Intrinsic fluctuations in sustained attention and distractor processing. The Journal of Neuroscience. 2014; 34(5):1724–1730. [PubMed: 24478354]

Galvan A, Hare TA, Parra CE, Penn J, Voss H, Glover G, et al. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. The Journal of Neuroscience. 2006; 26(25):6885–6892. [PubMed: 16793895]

Germine LT, Duchaine B, Nakayama K. Where cognitive development and aging meet: Face learning ability peaks after age 30. Cognition. 2011; 118(2):201–210. [PubMed: 21130422]

Germine LT, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. Psychonomic Bulletin & Review. 2012; 19(5):847–857. [PubMed: 22829343]

Halberda J, Ly R, Wilmer JB, Naiman DQ, Germine LT. Number sense across the lifespan as revealed by a massive internet-based sample. Proceedings of the National Academy of Sciences. 2012; 109(28):11116–11120.

Hartshorne JK, Germine LT. When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the lifespan. Psychological Science. in press.

Hedden T, Gabrieli JDE. Insights into the ageing mind: a view from cognitive neuroscience. Nature Reviews: Neuroscience. 2004; 5(2):87–96. [PubMed: 14735112]

Kalechstein AD, Newton TF, van Gorp WG. Neurocognitive Functioning is Associated ith Employment Status: A Quantitative Review. Journal of Clinical and Experimental Neuropsychology. 2003; 25(8):1186–1191. [PubMed: 14566590]

Lam CM, Beale IL. Relations Among Sustained Attention, Reading Performance, and Teachers' Ratings of Behavior Problems. Remedial and Special Education. 1991; 12(2):40–47.

McAvinue LP, Habekost T, Johnson K, Kyllingsbæk S, Vangkilde S, Bundesen C, et al. Sustained attention, attentional selectivity, and attentional capacity across the lifespan. Attention, Perception, & Psychophysics. 2012; 74:1570–1582.

McGraw KO, Tew MD, Williams JE. The Integrity of Web-Delivered Experiments: Can You Trust the Data? Psychological Science. 2000; 11(6):502–506. [PubMed: 11202497]

Owsley C. Aging and vision. Vision Research. 2011; 51:1610–1622. [PubMed: 20974168]

Park M, Hood MM, Shah RC, Fogg LF, Wyatt JK. Sleepiness, parkinsonian features and sustained attention in mild Alzheimer's disease. Age and Ageing. 2012; 41(6):765–770. [PubMed: 22743151]

Ratcliff R, Thapar A, McKoon G. The effects of aging on reaction time in a signal detection task. Psychology and Aging. 2001; 16(2):323–341. [PubMed: 11405319]

Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J. `Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia. 1997; 35(6):747–758. [PubMed: 9204482]

Rosenberg M, Noonan SK, DeGutis J, Esterman M. Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. Attention, Perception, & Psychophysics. 2013; 75(3):426–439.

Sarter M, Givens B, Bruno JP. The cognitive neuroscience of sustained attention: where top-down meets bottom-up. Brain Research Reviews. 2001; 35:146–160. [PubMed: 11336780]

Schmidt EA, Schrauf M, Simon M, Fritzsche M, Buchner A, Kincses WE. Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. Accident Analysis & Prevention. 2009; 41:1087–1093. [PubMed: 19664450]

Silver H, Feldman P. Evidence for sustained attention and working memory in schizophrenia sharing a common mechanism. The Journal of Neuropsychiatry and Clinical Neurosciences. 2005; 17(3): 391–398. [PubMed: 16179663]

Staub B, Doignon-Camus N, Bacon E, Bonnefond A. Investigating sustained attention ability in the elderly by using two different approaches: Inhibiting ongoing behavior versus responding on rare occasions. Acta Psychologica. 2014; 146(1):51–57. [PubMed: 24378237]

Staub B, Doignon-Camus N, Després O, Bonnefond A. Sustained attention in the elderly: What do we know and what does it tell us about cognitive aging? Ageing Research Reviews. 2013; 12:459–468. [PubMed: 23261761]

Williams BR, Hultsch DF, Strauss EH, Hunter MA, Tannock R. Inconsistency in reaction time across the life span. Neuropsychology. 2005; 19(1):88–96. [PubMed: 15656766]

Yeatman JD, Wandell BA, Mezer AA. Lifespan maturation and degeneration of human brain white matter. Nature Communications. 2014; 5:4932.
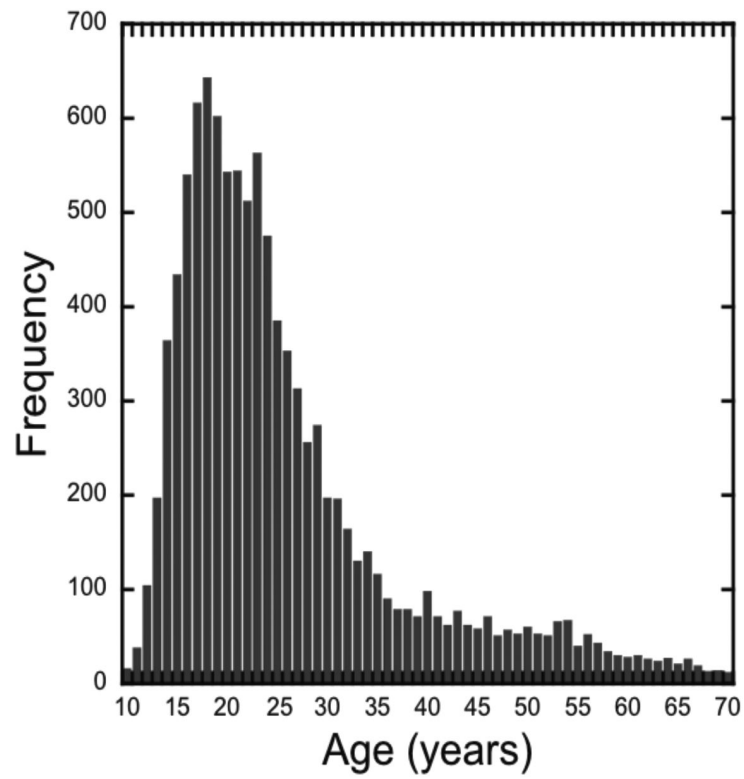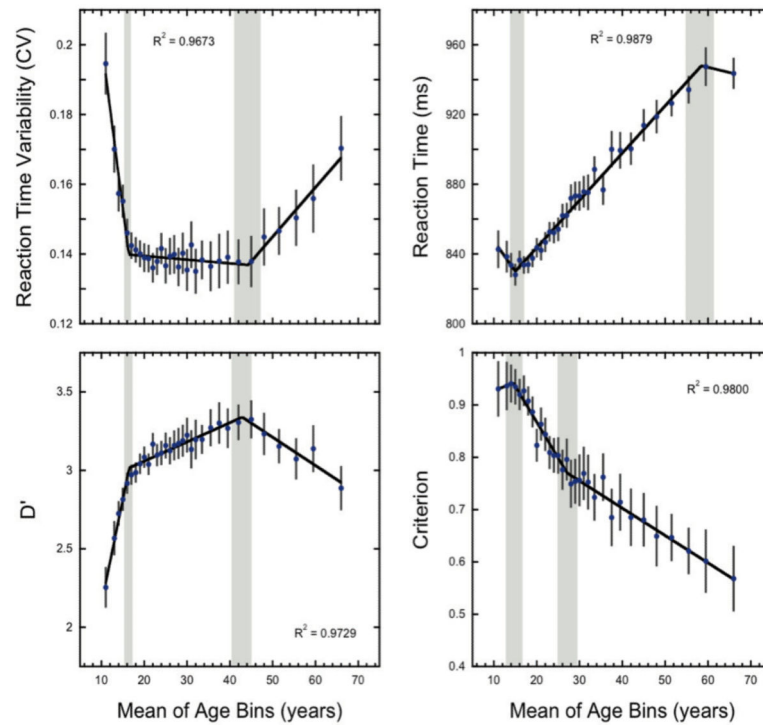
Figure 1.
Histogram of the number of participants by age.
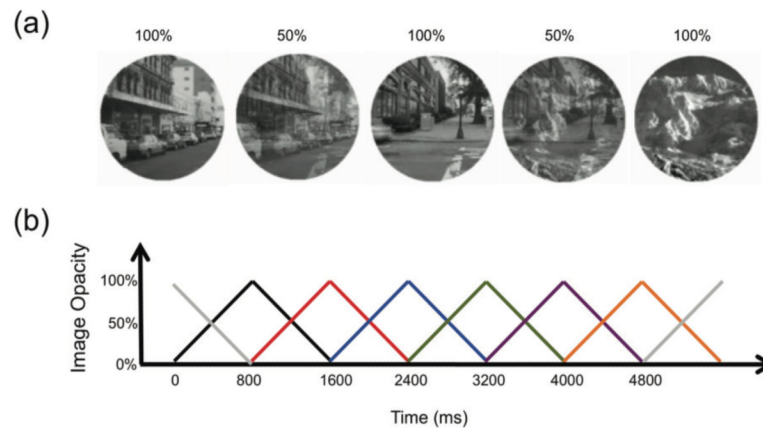
Figure 2.
Illustrations of the gradual continuous performance task (gradCPT). (a) Example images of scenes shown over two full trials with full and mixed opacity for three images. (b) Graph illustrating the linear transition in image opacity from one image to the next. Each colored triangle represents the opacity transition for a single image over time.
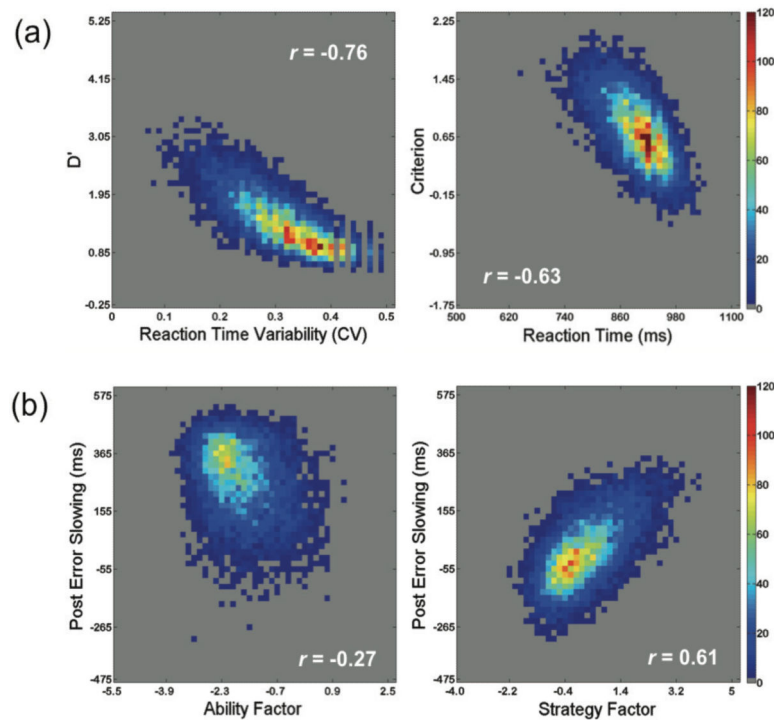
Figure 3.
Changes in sustained attention performance as a function of age. Lifespan trends in performance measures: variability of reaction time, mean reaction time, d', and criterion. Error bars show 95% confidence intervals. Black lines show the best fitting functions from the hierarchical regression analysis. Grey bars show the 95% confidence intervals for the estimated breakpoints. Reaction time variability is defined as the coefficient of variation (CV), or the standard deviation divided by the mean reaction time.
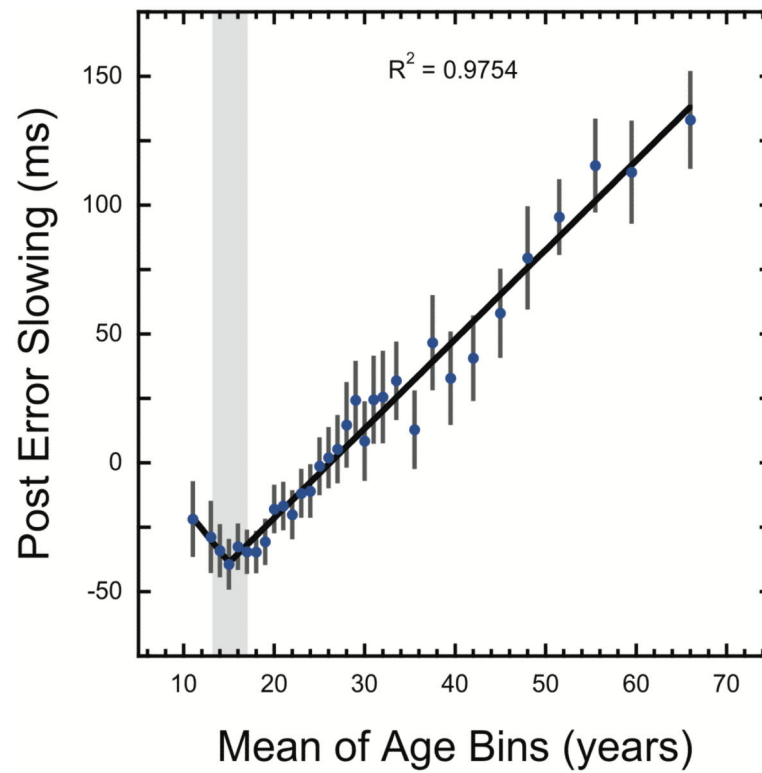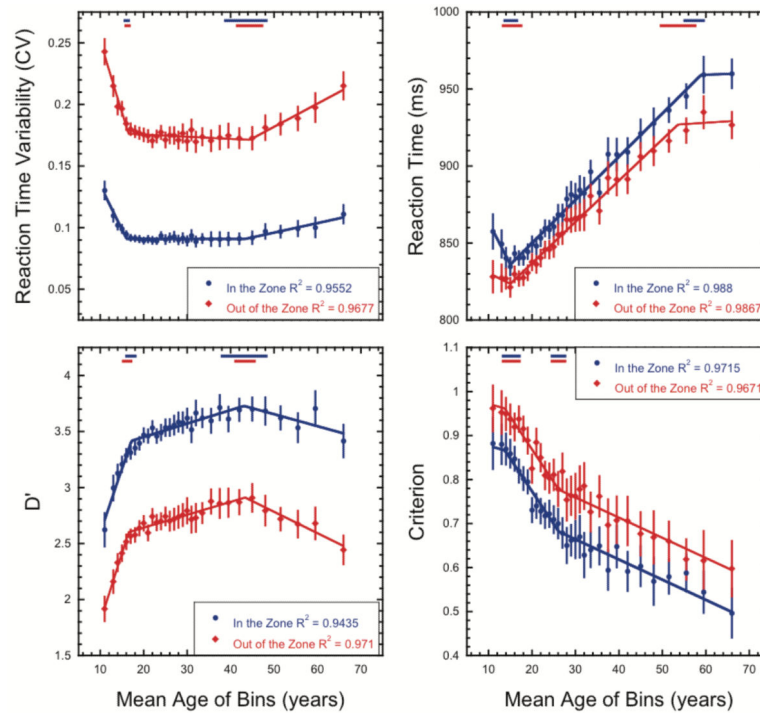
Figure 4.
Post-error slowing as a function of age. Error bars show 95% confidence intervals and the best fitting 2-phase segmented linear function is shown in black. The grey bar shows the 95% confidence interval for the estimated breakpoint.

Figure 5.
Correlation across sustained attention ability and strategy measures. (a) Density scatter plots showing significant correlations between measures related to ability (left panel) and strategy (right panel) across all participants. (b) Density scatter plots illustrating the stronger relationship between post error slowing and the calculated factor score related to strategy (right panel) than the calculated factor score related to ability (left panel). For all plots the color indicates the number of participants represented at a given location.

Figure 6.
Changes in sustained attention performance as a function of age and zone. Lifespan trends in performance measures: reaction time variability, mean reaction time, d', and criterion. Performance when participants are "in the zone" is shown as blue circles and performance when participants are "out of the zone" is shown as red diamonds. Error bars show 95% confidence intervals. Blue and red lines show the best fitting functions from the hierarchical regression analysis for each condition. Red and blue dashes at the top of the graphs show the 95% confidence intervals for the estimated breakpoints.

Table 1

Statistical results of hierarchical regression used for model selection

	Straight Line vs. 1-Break	1-Break vs. 2-Break
Reaction Time	$F(2,28) = 7.08, p = 0.003$	$F(2,26) = 12.15, p = 0.0002$
Reaction Time Variability	$F(2,28) = 84.17, p < 0.0001$	$F(2,26) = 43.63, p < 0.0001$
D'	$F(2,28) = 28.03, p < 0.0001$	$F(2,26) = 109.8, p < 0.0001$
Criterion	$F(2,28) = 25.13, p < 0.0001$	$F(2,26) = 4.29, p = 0.025$
Post-Error Slowing	$F(2,28) = 25.13, p < 0.0001$	$F(2,26) = 0.37, p = 0.69$

Table 2

Results of Factor Analysis

	All		10 – 16 years		17 – 45 years		46 – 70 years		All	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
RT	.115	.909	.197	−.883	.091	.910	.201	.921	.203	.901
RT Variability	−.938	.018	−.937	.042	−.934	.010	−.958	−.005	−.937	.046
D'	.935	.017	.931	.040	.931	.008	.939	−.003	.928	.035
Criterion	.121	−.895	.185	.899	.097	−.891	.282	−.812	.025	−.847
Post-Error Slowing									−.222	.785

Table 3

Parameter estimates (95% confidence intervals) for performance measures for "In the Zone" and "Out of the Zone" conditions. Parameters X0 and X1 are the 1st and 2nd breakpoints, respectively.

		Reaction Time	Reaction Time Variability	D'	Criterion
In the Zone	**Intercept**	919.6 (867.5 to 971.7)	0.2052 (0.191 to 0.220)	1.452 (1.1 to 1.8)	0.9047 (0.6 to 1.1)
	Slope1	−5.574 (−9.7 to −1.5)	−0.007112 (−0.008 to −0.006)	0.1144 (0.09 to 0.14)	−0.002786 (−0.022 to −0.017)
	X0	15.0 (13.7 to 16.4)	16.08 (15.6 to 16.6)	17.18 (16.2 to 18.2)	13.73 (13.0 to 17.2)
	Slope2	2.805 (2.64 to 3.0)	−0.0000004 (−0.0001 to 0.0001)	0.01202 (0.008 to 0.016)	−0.01487 (−0.02 to −0.01)
	X1	58.93 (54.9 to 59.5)	43.36 (38.6 to 48.1)	43.1 (38.0 to 48.2)	26.15 (24.4 to 27.9)
	Slope3	0.0948 (−1.9 to 2.1)	0.00077 (0.0005 to 0.0010)	−0.01071 (−0.017 to −0.004)	−0.004587 (−0.005 to −0.004)
Out of the Zone	**Intercept**	844.6 (801.2 to 888.1)	0.3669 (0.343 to 0.391)	0.5862 (0.29 to 0.88)	1.0 (0.8 to 1.2)
	Slope1	−1.404 (−4.7 to 1.9)	−0.01156 (−0.013 to −0.0099)	0.1218 (0.10 to 0.14)	−0.002786 (−0.02 to 0.02)
	X0	15.0 (13.0 to 17.2)	16.5 (15.96 to 17.03)	16.55 (15.9 to 17.2)	13.73 (13.0 to 17.2)
	Slope2	2.68 (2.5 to 2.9)	−0.0001743 (−0.0004 to 0.00002)	0.01152 (0.009 to 0.014)	−0.01487 (−0.02 to −.01)
	X1	53.59 (49.4 to 57.8)	44.41 (41.6 to 47.3)	43.36 (40.9 to 45.8)	26.15 (24.4 to 27.9)
	Slope3	0.1821 (−1.0 to 1.4)	0.001886 (0.0015 to 0.0023)	−0.01911 (−0.02 to −0.01)	−0.004587 (−0.005 to −0.004)