+Model
ARTICLE IN PRESS

Journal of Applied Research in Memory and Cognition xxx (2020) xxx–xxx

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control☆

Jessie Martin*, Cody A. Mashburn, Randall W. Engle

Georgia Institute of Technology, United States

We evaluated the predictive value of the Armed Services Vocational Aptitude Battery (ASVAB) at the latent level, using multitasking as a proxy for real-world job performance. We also examined whether adding measures of attention control to the ASVAB could improve its predictive validity. To answer these questions, data were collected from 171 young adults recruited from the Georgia Institute of Technology and the greater Atlanta community. Both regression and latent variable analyses revealed that the ASVAB does predict multitasking at the latent level but that measures of attention control add substantial predictive validity in explaining multitasking above and beyond the ASVAB, fluid intelligence, and processing speed. Theoretical as well as practical applications of these results are discussed in terms of theories of attention control, and potential cost savings in selection for military positions.

General Audience Summary

This study served as an in-house test to see if we could improve the prediction of the Armed Services Vocational Aptitude Battery (ASVAB). A total of 171 individuals came into the lab, completed a series of practice tests (due to the propriety nature of the ASVAB), some basic measures of cognitive ability, and measures focused on the ability to control ones attention. We then measured the degree to which these abilities predicted performance on a complex series of computer-based tasks designed to approximate real-world job performance.

Our results showed that while the ASVAB does a good job of predicting complex task performance, prediction rates are improved when measures of attention control are included as well. From this data set we have selected a series of tasks that are easy to administer and understand which we suggest be included in testing active duty personnel.

Keywords: ASVAB, Attention control, Multitasking, Fluid intelligence

The United States military has a vested interest in maximizing the validity of its selection instruments. One such instrument is the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB is a standardized test administered to all individuals enlisting in the United States military. The test and its derivations (e.g., the Armed Forces Qualification Test) are used to screen and select personnel for jobs across all branches of the military and were designed with these utilitarian goals in mind. Although performance on the ASVAB approximates some more psychologically informed tests, including the SAT (Frey & Detterman, 2004), the ASVAB was constructed relatively independently of psychological theory surrounding general cognitive abilities and intelligence testing (Roberts et al., 2000). As such, we believe that the addition of metrics grounded in psychological theory

Author Note.

Jessie Martin, Cody A. Mashburn, & Randall W. Engle, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30313, United States.

This work was supported by a grant from the Office of Naval Research (N00014-12-1-1011) to Randall Engle.

* Correspondence concerning this article should be addressed to Jessie Martin, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30313, United States. Contact: jessie.martin@gatech.edu.

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. *Journal of Applied Research in Memory and Cognition* (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

may be beneficial in improving its prediction. We conducted this study at the request of the Office of Naval Research as a lab-based study to see whether the validity of the ASVAB for predicting success in military vocational training programs could be improved by increasing the test's emphasis on fluid abilities.

According to Ree and Carretta (1994), the correlation between ASVAB scores and job performance is approximately .40, meaning that the ASVAB accounts for around 16% of the overall variability, leaving 84% of the variance in job performance unaccounted for. Additionally, although it was intended to measure aptitude, that is, the *propensity* to do well, research has shown that the ASVAB primarily reflects acculturated or acquired knowledge, or crystallized intelligence (Gc). Roberts et al. (2000) submitted Naval recruits' ASVAB scores along with numerous measures thought to reflect Gc, fluid intelligence (Gf), a person's ability to perform novel reasoning, and general speed to an exploratory factor analysis. Of the eight subtests still reflected on the current ASVAB, five loaded substantially onto a Gc factor and another three loaded substantially onto a factor reflecting technical knowledge (which was correlated with Gc). Three of the eight subtests also loaded onto a Gf factor, but each shared substantial variance with either the Gc or technical knowledge factor. Thus, no ASVAB subtest uniquely and substantially measured Gf.[1] The test therefore does little to measure the ability to do novel problem solving, reasoning, or complex learning (Roberts et al., 2000). In short, the ASVAB is primarily an assessment of existing knowledge, not one of aptitude or ability to acquire new knowledge. Thus, Roberts et al. (2000) suggested increasing the ASVAB's emphasis on Gf to broaden its predictive validity.

While Gc and Gf tend to be correlated (Kvist & Gustafsson, 2008), theorizing around the Gc-Gf relationship provides a basis for decoupling them. Cattell's (1987) investment theory posits that Gc is the result of exposure to sources of information *and* the individual's Gf at the time of learning; high Gf along with high investment by the individual results in high Gc. By the logic of investment theory, there may be many individuals who could successfully perform complex jobs (i.e., who have the necessary Gf for effectively learning new skills) but perform poorly on tests emphasizing culturally derived knowledge.

The shifting nature of military work also compounds the need to increase the ASVAB's emphasis on fluid abilities. Continual technological development necessitates recurrent learning and adaptation on the part of military technicians, a demand for which individuals with higher Gf are better suited. As vocational demands become more complex and change over time, rapid problem-solving and adaptation to novel situations will become increasingly valued traits among military recruits (National Research Council, 1999). We thus agree with Roberts et al.'s (2000) conclusion that supplementing the existing ASVAB with additional measures of fluid ability, including attention control.

[1] For a seminal paper describing and differentiating Gf and Gc, see Cattell (1963).

Attention Control and Its Measurement

Attention control is the ability to focus attention, regulate the intensity of attention, and to resist or quickly recover from its capture by internally generated events, such as mind wandering, or externally generated events, such as environmental distractions. Individual differences in attention control are at least partly responsible for the relationship between higher order cognitive constructs, including working memory capacity and fluid intelligence (Engle, 2002; Kane et al., 2004; Shipstead, Harrison, & Engle, 2016). Further, differences in working memory capacity and attention control are predictive of an individual's ability to learn and perform complex tasks in a huge range of real-world situations (Kleider-Offut, Clevinger, & Bond, 2016; Morrow et al., 2003; Shute, 1991; Wood, Hartley, Furley, & Wilson, 2016).

Measures of attention control emphasizing vigilance have also proven to be important predictors for performance in a variety of military contexts (Shaw et al., 2010; Matthews, Warm, Shaw, & Finomore, 2014). Additionally, attention control measures are generally quite simple with little memory load (e.g., Roberts, Hager, & Heron, 1994). This simplicity means that they are relatively process-pure (i.e., they are less multiply determined than many other cognitive tasks) and are likely to predict performance across a wide range of situations, since broader abilities are likely to subsume them. It also means that they are easy to administer, and require little in the way of culturally derived knowledge to perform, suggesting that they may be an avenue to more equitable selection from culturally heterogeneous populations. These features make them ideal candidates for inclusion in a military selection battery.

One final consideration when designing and implementing measures of attention control for administration in applied settings is task design. We and others argue that attention control is fundamental to individual differences in higher-order cognition (Engle, 2018; Friedman & Miyake, 2017; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). However, many common paradigms used to study attention control are optimized for capturing within-subject variance (attributable to experimental manipulations) rather than between-subject variability (attributable to stable individual differences).[2] The former typically rely on reaction time difference scores[3] as their dependent variables, which Draheim, Mashburn, Martin, and Engle (2019) argue are psychometrically unsound for individual differences research. Instead, Draheim et al. (2019) advocated that those interested in individual differences in attention control adopt non-subtractive accuracy-based measures. As a test of these psychometric improvements, we included both reaction-time

[2] See Cronbach (1957) for a lengthier discussion of the differences between the experimental and individual differences approach.

[3] Reaction-time difference scores refer to scores derived from two theoretically related tasks. Typified by the calculation of interference effects, this often involves subtracting reaction times in an experimental condition in which stimuli and responses are consistent with one another (i.e., congruent trials) from a condition where participants must resolve conflict between two stimuli and/or between a stimulus and a response to respond appropriately (i.e., incongruent trials).

difference scores and accuracy-based measures of attention control in the present study.

The Present Study

We sought to test these assumptions with a preliminary lab-based study as a proof-of-concept to be followed by a large-scale study of Navy recruits. We elected to use multitasking ability as a proxy for work performance. Although measures of multitasking ability generally lack face validity for military vocational training programs, the need to manage multiple concurrent goals is an increasingly common feature of modern military work (National Research Council, 1999). Thus, multitasking is an increasingly common laboratory analogue for work performance (Barron & Rose, 2017; Colom, Martínez-Molina, Shih, & Santacreu, 2010; Hambrick, Oswald, Darowski, Rench, & Brou, 2010; Hambrick, Rench, et al., 2011; Redick, 2016).

The ASVAB, measures of Gf, and measures of attention control have all been associated with multitasking in prior research. ASVAB scores have been positively associated with multitasking in samples of Navy enlistees (Hambrick, Oswald, et al., 2010; Hambrick, Rench, et al., 2011). Redick et al. (2016) found working memory capacity was substantially related to multitasking at the latent level, but the relationship was completely explained by a combination of attention control and static memory storage capacity. Gf was also related to multitasking ability, but this association was not fully decomposable. This suggests unique contributions of Gf to multitasking. The present study is the first to examine the degree to which fluid ability measures add incremental validity to the ASVAB for the prediction of multitasking using latent variable analysis. This method is advantageous because it permits purer measurement of the constructs of interest, which is useful in assessing the relative importance of different cognitive mechanisms for predicting complex behavior.

Additionally, the present study included measures of processing speed. Processing speed has been central to psychometric theories of intelligence (Vernon, 1983), and has been proposed as a theoretical alternative to attention control (Fry & Hale, 1996; Rey-Mermet, Gade, Souza, Bastian, & Oberauer, 2019). However, other work suggests that processing speed measures may in fact reflect variance related to attention control (e.g., Cepeda, Blackwell, & Munakata, 2013). Including the processing speed measures allows us to compare the predictive capabilities of attention control and processing speed, as well as to document and characterize redundancies between latent variables underlying putative attention control and processing speed tasks.

In order to test the relationships between the constructs described, we conducted a series of sequential regression and structural equation models, each time with multitasking, our proxy for work performance, serving as the criterion. We approached these analyses with several a priori hypotheses:

Will fluid abilities (attention control and Gf) account for variance in multitasking beyond that accounted for by the ASVAB? If there is no benefit to prediction when fluid measures are added, they will not show incremental validity in regression analyses (H1).

Will non-subtractive accuracy-based attention tasks have superior prediction to reaction-time difference score-based measures? This hypothesis is based on preliminary evidence by Draheim, Tsukahara, Martin, Mashburn, & Engle (in press), which suggests that accuracy-based measures of attention control help create a more robust measure of attention control (H2).

Will attention control account for unique variance in multitasking above and beyond Gf? This hypothesis is speculative and based on the idea that Gf may not fully encompass domain-general fluid abilities (H3).

Will processing speed predict multitasking independent of our measures of attention control? This served as a source of discriminant validity, but was also included due to the elimination of a processing-speed measure on previous versions of the ASVAB (H4).

Method

Participants

The sample consisted of a subset of participants who were recruited from a related study on attention control (see Draheim et al., in press). These individuals completed an additional two sessions consisting of our ASVAB, multitasking, and processing speed measures. All other measures were collected in the four sessions prior. The initial sample contained 183 participants sampled from the Georgia Institute of Technology and greater Atlanta communities. A total of eight individuals were excluded due to attrition. An additional four subjects were excluded because of extreme values leading to unacceptably high skewness and kurtosis on one of the multitasking variables, yielding a final sample of 171. Any additional missing data were imputed via maximum likelihood estimation. All were between the ages of 18 and 35 and spoke English fluently. This study was approved by the Georgia Institute of Technology IRB under protocol #H18146.

Measures

Armed Services Vocational Aptitude Battery practice test (Powers, 2011). Given that the ASVAB is a restricted and proprietary selection tool, we were not able to administer the current version of the test to participants. Instead, participants took a test comprising questions from an ASVAB practice test booklet (Powers, 2011). Each subtest was chosen from among four alternatives by the researchers, the main criterion for selection being to select the version of the subtest with the broadest content. This was mainly a concern for the General Science subtest, some versions of which were comprised largely of meteorological questions. For other subtests where specific sub-domain knowledge was less of a concern, such as the Word Knowledge or Assembling Objects subtests, the choice of form was random. Due to time constraints, we could not give participants every question from each subtest without the risk of rushing

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. *Journal of Applied Research in Memory and Cognition* (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

Table 1
Timing Information for the ASVAB Practice Test

Subtest	Number of questions	Time allowed (s)
General Science (GSScore)	20	528
Arithmetic Reasoning (ARScore)	24	1728
Word Knowledge (WKScore)	28	528
Paragraph Comprehension (PCScore)	12	624
Mathematics Knowledge (MKScore)	20	1152
Electronics Information (EIScore)	16	432
Auto and Shop Information (ASISCore)	20	528
Mechanical Comprehension (MCScore)	20	912
Assembling Objects (AOScore)	20	720

participants and deviating from standard timed ASVAB administration procedures. To solve this problem, one third of the items on each subtest were not tested. Items included were randomly selected by the researchers using a random number generator and the resulting items were administered to all participants. The test was administered via computer, maintaining the timing integrity per question of the sample materials, and participants registered their response via keys labeled A, B, C, and D on a standard keyboard. Each subtest was timed such that each participant had no longer than two hours to complete the entire test. There were optional untimed breaks between subtests. Specific information about the number of questions and timing for each subtest is reported in Table 1. For purposes of the regression analyses reported below, the dependent variable was the total number of items correct across subtests.

Attention control. Real-time demonstrations of all attention control tasks listed below can be found on our website: http://englelab.gatech.edu/taskdemos.html.

Antisaccade (Hallett, 1978; Hutchison, 2007; Kane, Bleckley, Conway, & Engle, 2001). Participants saw a fixation cross lasting a random amount of time between 2000 and 3000 ms followed by an alerting tone for 300 ms. After the alerting tone, a distractor (a flashing asterisk) appeared for 300 ms on either the left or the right of the screen approximately 12.3° visual angle from fixation followed immediately by a target "Q" or an "O" for 100 ms 12.3° from fixation on the *opposite* side of the screen from the asterisk. The location of the asterisk and target letter were both masked for 500 ms by "##". The participant was instructed to ignore the asterisk and instead look away to the other side of the screen to catch the target "Q" or "O". Participants had as much time as needed to respond to which letter appeared by pressing the associated key on the keyboard. Participants completed 72 trials, with trial-by-trial feedback for 500 ms following each response, and then a 1000 ms waiting period until the fixation cross appeared again to indicate a new trial was beginning. The dependent variable was the number of correctly identified target letters.

Arrow Flanker Task (Eriksen & Eriksen, 1974). The participant was shown a series of arrows in one of three configurations and reported which direction the middle arrow was pointing (see Figure 2). The subject was asked to indicate direction the central arrow was pointing by pressing the "z" (left) or "/" (right) key. These keys had the words LEFT and RIGHT taped onto them to assist with response mapping. A total of 144 trials were administered; 96 congruent and 48 incongruent, with a randomized 400–700 ms intertrial interval. The dependent variable was the flanker interference effect and was calculated by subtracting each participant's mean reaction time on congruent trials from their mean reaction time on incongruent trials, excluding inaccurate trials.

Color Stroop Task (Stroop, 1935). Participants were shown words printed in various colors that fall into different categories. The words were either congruent with the color (e.g., the word RED printed in red ink) or incongruent with the color (e.g., The word BLUE printed in green ink). The participant's task was simply to indicate which color the word was printed in by pressing the 1, 2, or 3 key on the number pad. To assist with response mapping, the keys had a piece of paper of the corresponding color taped onto them. A total of 144 trials were administered, 96 congruent and 48 incongruent, with a randomized 400–700 ms intertrial interval and a 5000 ms response deadline. The dependent variable was the Stroop interference effect and was calculated by subtracting each participant's mean reaction time on congruent trials from their mean reaction time on incongruent trials, excluding inaccurate trials.

Selective Visual Arrays (Martin et al., under revision; Shipstead, Lindsey, Marshall, & Engle, 2014; Vogel, McCollough, & Machizawa, 2005).[4] Participants saw a target array of blue and red rectangles differing in orientation. Prior to each trial, the participant was cued to attend to either the red or blue rectangles. Next, after a 900 ms delay, the target array was presented for 250 ms. Finally, the test array was presented which contained only the rectangles of the cued color, one of which was highlighted by a white dot. The rectangle with the white dot was in a different orientation from the target array on 50% of the trials. The participant was then asked whether or not the rectangle had changed orientation from the initial presentation. Each array contained either 5 or 7 rectangles per color (10 and 14 total), and a total of 48 trials were presented for each array size. The dependent variable was a capacity score (k), which is calculated using the single probe correction (see Cowan et al., 2005; Shipstead et al., 2014). This calculation is $N \times$ (% hits + % correction rejections − 1), where N is the set size for that array. This calculation results in two separate k scores, one for set size 5 and one for set size 7, and the final dependent variable was the average k for these two set sizes.

Sustained Attention to Cue Task (SACT; Draheim et al., in press). This task was designed as an accuracy-based analogue of the psychomotor vigilance task (PVT). See http://englelab.gatech.edu/taskdemos for a live speed demonstration. In this task participants needed to sustain their attention on a visual circle cue and identify a target letter presented briefly at the center of the cue. The stimuli for the task were presented against a grey background. Each trial started with a central black

[4] Although this task is not traditionally included in attention control batteries, evidence strongly suggests that is an attention control task (Martin et al., submitted for publication; Vogel et al., 2005).

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. *Journal of Applied Research in Memory and Cognition* (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

fixation. On half of the trials, the fixation was presented for 2 s and for the other half the fixation was presented for 3 s. After the fixation, there was a 300 ms tone, then a large white circle cue was presented in a randomly determined location on either the left or right side of the screen. To orient the participant on the circle cue, the large circle began to immediately shrink in size until it reached a fixed size (1.5 s). Once the cue reached the fixed size, there was a variable wait time (equally distributed amongst 2, 4, 8, and 12 s) and then a white asterisk designed to serve as a distractor appeared at the center of the screen. The asterisk blinked on and off in 100 ms intervals for a total duration of 400 ms. A 3×3 array of letters was then displayed at the center of the circle cue. The letters in the array consisted of B, D, P, and R. The central letter was the target letter and was presented in a dark grey font. The non-target letters were presented in black font with each letter occurring twice. After 125 ms the central letter in the array was masked with a # for 1000 ms. Only the central target letter was masked. After the mask, the response options were displayed in boxes horizontally across the upper half of the screen. The participant used the mouse to select whether the target was a B, D, P, or R. Feedback was given during the practice trials but not the experimental trials. Accuracy rate was the dependent variable.

Fluid Intelligence. Raven's Advanced Progressive Matrices–Odd problems (RAPM; Raven & Court, 1998). Participants were shown abstract shapes in a 3×3 matrix. The shape in the bottom-right was missing, and the participant had to select which of the eight answer options best completed the overall pattern by clicking that option on the screen. Participants had 10 min to complete 18 problems. The number of correct responses was the dependent variable.

Letter Sets (Ekstrom, French, Harman, & Dermen, 1976). On each problem, the participant was presented five sets of letters, each containing four letters that follow a particular rule. Instructions were to find the rule that applied to four of the five letter sets, and then indicate the set that violates the rule by clicking that set on the screen. Participants had 7 min to complete 30 problems. The dependent variable was the number of correct responses.

Number Series (Thurstone, 1938). Participants were presented a sequence of numbers and needed to identify the response option that was the next logical number in the sequence by clicking the correct number from five total response options. Participants had 5 min to answer 15 problems, with the number of correct responses serving as the dependent variable.

Processing Speed. All processing speed measures were computerized versions of paper and pencil tests. In each case, participants were instructed to respond as quickly and accurately as possible, but consistent with standard administration procedures, were not alerted to the time limits of each task in the instruction phase.

Digit String Comparison (Redick, Unsworth, Kelly, & Engle, 2012; Salthouse & Babcock, 1991). In this version of the Digit String Comparison Task, participants viewed strings of three, six, or nine digits appearing to the left and right of a central line. The digit strings could either be the same or differ by a single digit. If different, the mismatching digit could appear in any location in the string. Participants indicated their response by clicking on a button on the screen labeled SAME for identical strings or DIFF for mismatching strings. Digits were printed in white size 18 Courier New font on a black background. After completing six practice trials, participants completed as many trials as possible within two 30-s blocks of the task. The dependent variable was the number of accurate responses across both blocks.

Letter String Comparison (Redick et al., 2012). The Letter String Comparison task was identical to the Digit String Comparison, except that participants viewed and made judgements about strings containing three, six, or nine Latin characters.

Digit Symbol Substitution (Wechsler, 1997). This adaptation of the Digit Symbol Substitution Task was modified to make it more amenable to computer administration and response collection via a standard number pad. The symbols used were the same as the paper and pencil version of the task, and we endeavored to maintain the same basic demands. However, rather than viewing digits and reporting corresponding symbols, this task required participants to view symbols and report the corresponding digits. On each trial, participants were presented with two boxes one atop the other in the center of the screen. A symbol appeared in the bottom box. Participants were to indicate via key press with the digit that belonged in the top box. They could consult a table presented at the top of the screen or rely on their memory of the table. After ten practice trials, participants completed 90 sof the task. The dependent variable was the number of correctly reported digits during that 90-s period.

Multitasking. For all tasks, the dependent variables were the sum score of all components of each task. Regression analyses used a composite score while z-score transformations were used for all structural equation models due to the inconsistent scaling across task dimensions. Please see supplemental materials for figures.

Control Tower (Redick et al., 2016). The control tower task was designed with one primary task that was completed while also managing interruptions from four other tasks. Both visual and auditory processing was necessary for task success. Participants were presented with a visual display with various subtasks to complete (see Figure A in the supplemental materials). For the primary task, participants searched through an array of numbers, letters, and symbols on the left side of the screen and selected the appropriate items from the array on the right side of the screen, as indicated by task instructions: for numbers, participants clicked on the identical numbers in the right array; for letters, participants clicked on the letter that preceded it alphabetically in the right array; and for symbols, participants referred to a consistently mapped symbol *code book* and clicked on the relevant symbols in the right array.

Participants were instructed to complete as many array comparisons as possible over the duration of the task. Meanwhile, distractor tasks (radar, airplane, color, problem-solving) were presented visually and via headphones to reference items listed below the arrays at specific times or when specific colors were presented. For the radar task, participants were instructed to click either the *inside* or *outside* button below the radar when a blip occurred. For the airplane task, requests for landing on one of three runways were presented via headphones and the participant either agreed to or denied the request, according to the availability listed in the button marked *runway*. For the color task, a color flashed briefly and the participant pressed one of three "error" buttons according to the mapping presented in the button marked *protocol*. For the problem-solving task, trivia, logic, and general knowledge questions were presented via headphones, and participants clicked on one of three possible answers presented at the bottom of the screen. Participants completed one, 10-min block of the task. The participant's primary score was the number of correct number, letter, and symbol comparisons completed (no theoretical maximum score). The participant's distractor score was the sum of the correct decisions across the various distractor subtasks (out of 30). Both the primary and distractor scores were used as dependent variables.

***SynWin* (*Elsmore, 1994*; *Hambrick, Oswald, et al., 2010*).** SynWin is a proprietary multitask which requires simultaneous processing of four independent tasks across both auditory and visual modalities (see Figure B in the supplemental materials). Participants completed three 5-min blocks of the task. The participant's score for each block was determined by a formula that combined the points earned across all subtasks, and the dependent variable for our analyses was the mean score across the three blocks. This composite score was used as the dependent variable. The subtasks were as follows:

1. Probe-recognition: A six-letter list was presented for 5 s and then disappeared. For the remainder of the task, a probe letter was presented every 10 s. The participant indicated (via mouse click) whether the probe letter was one of the six letters presented on the list. Ten points were added to the participant's score for correct responses, and 10 points were subtracted for incorrect and missed responses.
2. Arithmetic: The participant mentally added two three-digit numbers and reported the correct sum via mouse click. The arithmetic subtask was entirely self-paced. Twenty points were added for correct responses and 10 points subtracted for incorrect responses.
3. Visual monitoring: The participant monitored the level on a gauge. As the gauge approached empty, participants clicked that section of the screen to "reset" it before it reached "empty." Points were awarded each time the participant reset the gauge, with more (up to 10), and 10 points are subtracted for every second that the gauge remained at zero before being reset.
4. Auditory monitoring: High (2000-Hz) and low (1000-Hz) frequency tones were presented every 10 s. The participant clicked a button within the quadrant when the rarely occurring high-frequency tone was presented (20% of all tones). Ten points were added for hits, and 10 points were subtracted for misses and false alarms.

***Foster Task*.** The Foster Task is a novel multitask modeled on SynWin. It too required participants to perform four concurrent subtasks (see figure c in the supplemental materials). Unlike SynWin, each of the Foster subtasks was visual in nature. The dependent variable was the aggregate score across all four subtasks. Participants increased their score by responding appropriately within a brief period, resulting in 100 points being added to the score for that subtask. Errors led to a subtraction of 100 points, while failing to respond within a brief response window resulted in a precipitous drop in points for that task until a response was made or the counter reaches zero. Participants completed three 5-min blocks of the task. Due to missing data from computer malfunctions, the dependent variable was the total score across the first administration. The subtasks were as follows:

1. Telling time: Participants saw an unlabeled analog clock face and were asked to choose, via mouse click, which of four alternatives corresponded to the position of the clock's hour and minute hand.
2. Visual monitoring: Participants were presented with a disk that would start spinning periodically. They were instructed to click on the disk as soon as possible to make it stop moving.
3. Word recall: Participants were prompted to remember a word printed in green. After a brief display, the word was removed. Shortly thereafter, a new prompt would appear in red asking whether this new word was the one they were supposed to remember previously. They responded either yes or no by mouse click and a new green to-be-remembered word was presented.
4. Math problems: Participants solved simple two term math problems. They were either addition, subtraction, multiplication, or division, and the participant chose the correct answer from among four possible answers.

Results

Descriptives

Raw descriptive statistics for all tasks are provided in Table 2 of the supplemental materials. Zero-order correlations can be found in Table 3. All data were standardized for inferential analyses. Full reliability, including test-retest can be found in Draheim et al. (2019) and Draheim et al. (in press).

Sequential Regression

Sequential regression was used to calculate how much incremental validity measures added in predicting multitasking performance over and above the ASVAB. Table 4 shows the amount of performance in multitasking that can be predicted from the ASVAB and various combinations of the attention control tasks. The criterion measure was a z-score composite multitasking score formed by taking the mean of each participant' standardized score across the four measures of multitasking ability. The column labeled "Predictor" lists the tasks in each model, the "Standardized B" column lists the standardized regression (beta) weight for each predictor which can be used to assess the strength of a predictor relative to other predic-

Table 4
Summary of Sequential Regression using the Different Attention Control Tasks

Criterion	Model	Predictor	Standardized *B*	Model R^2	ΔR^2
Multitask composite	1			**.467*****	–
		ASVAB	**.684*****		
	2			**.538*****	**.07*****
		ASVAB	**.579*****		
		Antisaccade	**.269*****		
		Flanker Effect	.044		
		Stroop Effect	−.091		
	3			**.573*****	**.036*****
		ASVAB	**.482*****		
		Antisaccade	**.19****		
		Flanker Effect	.051		
		Stroop Effect	−.078		
		Selective Visual Arrays	**.227*****		
		SACT	.036		
	4			**.567*****	−.006
		ASVAB	**.484*****		
		Antisaccade	**.196****		
		Selective Visual Arrays	**.227*****		
		SACT	.03		

Note. Bolded values are statistically significant. SACT = Sustained Attention to Cue Task.
**p* < .05. ***p* < .01. ****p* < .001

tors in the model, the "Model R^2" column lists the proportion of total variance accounted for by all predictors in the model combined, and the "ΔR^2" column lists the percent change in variance accounted for from one model to the next.

The first model (Model 1) showed that when the total score of the ASVAB alone was used to predict multitasking performance, it accounted for 46.7% of the variance in the multitasking criterion. The second model included the ASVAB as well as three of our attention control tasks (the antisaccade, the Stroop effect, and the flanker effects, the latter two of which are computed as reaction time difference scores). This set of attention control tasks has been commonly used in previous work to make an attention control latent factor. Model 2 accounted for 53.8% of the variation in our multitasking composite score, a statistically significant 7% increase, $F(3) = 8.41$, $p < .001$. Further, the beta-weights for each task show that this increase in prediction was based on the antisaccade task, as the contributions of the Stroop and the flanker effects were not significant.

Model 3 tested whether adding the non-subtractive accuracy-based attention control tasks to the model would improve model fit over and above the attention control tasks in Model 2. This would indicate that these tasks capture variance in multitasking not accounted for by the antisaccade, Stroop, and flanker, which is expected given the noted psychometric issues with the reaction time difference score tasks. Indeed, Model 3 accounted for 57.3% of the variation in multitasking, a statistically significant improvement over Model 2, $F(2) = 6.88$, $p = .001$. Finally, Model 4 tested whether the reaction-time difference score measures were necessary to achieve this level of prediction. Removing the Stroop and flanker effects via backwards elimination led to no reliable change in model fit from Model 3, $F(2) = 1.235$, $p = .293$. The Stroop and flanker effects were thus unnecessary, and all increased prediction of multitasking can be attributed to the accuracy-based attention control tasks.[5]

Next, we tested whether this accuracy-based attention control task set accounted for variance beyond Gf measures. Table 5 summarizes a sequential regression testing the amount of variance in multitasking predicted by the ASVAB, the ASVAB plus the Gf measures, and the ASVAB plus Gf measures and accuracy-based attention control tasks. Model 1 included the ASVAB alone, Model 2 included the ASVAB and all three Gf measures, and Model 3 included the ASVAB, the Gf measures, and the attention control measures. Once again, the ASVAB total accounted for 46% of performance in our multitasking composite criterion. Model 2, which added Gf measures, accounted for 55.5% of multitasking variance, improving model fit by 8.8%, $F(3) = 10.962$, $p < .001$. Finally, Model 3, which included the ASVAB, Gf, and accuracy-based attention control measures accounted for 62% of our multitasking composite score. Thus the attention control measures accounted for an additional 6.4% of multitasking performance above and beyond the ASVAB and measures of Gf, a statistically significant increase, $F(3) = 9.193$, $p < .001$. Both attention control and Gf led to significant improvements in model fit.

[5] While the SACT did not reliably increase model fit, we retain it in the model for the next set of analyses because we subsequently use it in forming attention control latent factors, as using at least three indicators of a latent variable is recommended practice (Kline, 2016). It is justifiable to retain the SACT for further analyses and not the Stroop and flanker effects because the SACT showed much stronger zero-order correlations with the other attention control tasks and multitasking than either the Stroop or flanker effects. This suggests that is likely redundant to the antisaccade and selective visual arrays.

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. Journal of Applied Research in Memory and Cognition (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

Table 5
Summary of Sequential Regression using Gf and Accuracy-Based Attention Control Tasks

Criterion	Model	Predictor	Standardized *B*	Model R^2	ΔR^2
Multitask composite	1			**.467*****	–
		ASVAB	**.684*****		
	2			**.555*****	**.088*****
		ASVAB	**.366*****		
		RAPM	**.177***		
		Letter Sets	**.242****		
		Number Series	.079		
	3			**.620*****	**.064*****
		ASVAB	**.259*****		
		RAPM	**.159***		
		Letter Sets	**.193****		
		Number Series	.057		
		Antisaccade	**.132***		
		Selective Visual Arrays	**.207*****		
		SACT	.043		

Note. Bolded values are statistically significant. RAPM = Raven's Advanced Progressive Matrices—Odd problems; SACT = Sustained Attention to Cue Task.
*$p < .05$. **$p < .01$. ***$p < .001$

Structural Equation Models

Structural equation models were used to clarify the theoretical relationships between the ASVAB, multitasking, attention control, fluid intelligence, and processing speed factors at the latent level. For all structural equation models, solid lines represent significant paths and dotted lines represent non-significant paths. The chi-square values, degrees of freedom, and chi-square significance are reported. The chi-square assesses overall fit of the reproduced variance-covariance matrix (derived by a theoretically specified model) to the actual variance-covariance matrix observed in the data; non-significant values suggest a well-fitting model. Although a non-significant chi-square value is preferred, it is very sensitive to sample size. As such, the chi-square value alone is not sufficient to accept or reject a model. Models must be considered in holistic terms based on multiple fit indices. The confirmatory fit index (CFI) and the Root Mean Square Error of Approximation (RMSEA) are also presented. The CFI compares model fit to a null model and is considered to be a good fit if the CFI is larger than .90. A CFI of over .90 indicates that the model of interest improves fit by 90% relative to a null model. The RMSEA is a parsimony adjusted fit index. Models with an RMSEA less than .08 are considered to be an acceptable fit, with an RMSEA of .06 or lower considered to be a good fit (Bollen & Long, 1993). All models described below had acceptable to good fit based on their constellation of fit indices.

Figure 1 shows the relationship between ASVAB total scores and a multitasking latent factor. First, multitasking showed a coherent latent factor suggesting it is an integrated ability important across all three multitasks. The ASVAB factor accounted for 77.4% of the variance ($.88^2$) in our multitask factor, replicating the strong relationship seen in the regression analyses. Next, we tested whether adding fluid intelligence and attention control factors provided independent prediction.

In Figure 2, the ASVAB uniquely accounted for 16% of the variance in performance in multitasking, and was the only significant predictor. However, the loadings of the Stroop and the flanker effects on the attention control factor were quite low, suggesting that the majority of the prediction between attention control and multitasking was based on the Antisaccade. A second model was run with the new accuracy-based attention control indicators to see if this increased attention controls prediction of multitasking ability.

In Figure 3, when the new accuracy-based attention control and fluid intelligence were included in a model with the ASVAB measures, and all indicators were allowed to correlate, only our attention control and fluid intelligence factors significantly predicted multitasking ability (uniquely accounting for 22.1% and 20.2% of the variance in multitasking, respectively). Thus, when accuracy-based attention control measures were used the ASVAB no longer predicted multitasking at the latent level. These results suggest that the lack of a relationship between attention control and multitasking in Figure 2 was due to the instability of the attention control latent factor, rather than a lack of a theoretical relationship between attention control and multitasking above and beyond the ASVAB. Further, the degree to which the ASVAB reflects differences in multitasking was likely due to its relationship to attention control.

Given that the non-subtractive accuracy based attention control measures employed in this study are novel, it is reasonable to wonder whether they do in fact measure the construct of interest. One possibility is that the shift into accuracy has made speed of processing information more relevant to these tasks. Earlier versions of the ASVAB included sections designed to measure processing speed (Roberts et al., 2000). In order to verify that we were not simply re-introducing this variance, we included computerized versions of traditional paper and pencil processing speed tasks in our test battery. This test of discriminant validity was run to ensure that our measure of attention control reflected the theoretical construct of interest. This analysis is summarized in Figure D of the supplemental materials. Processing speed did

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. *Journal of Applied Research in Memory and Cognition* (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

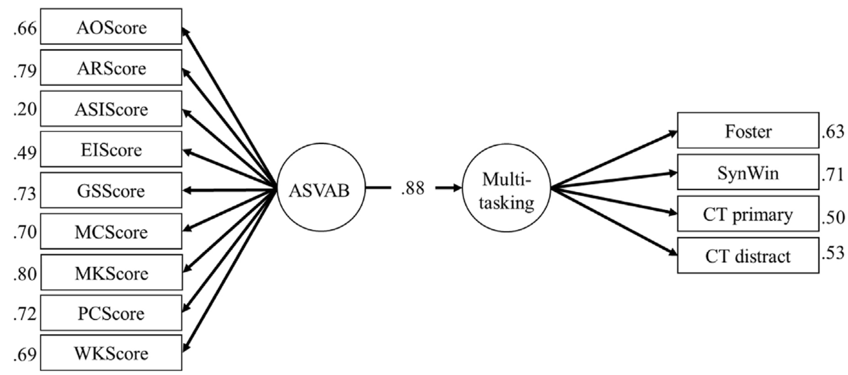

Figure 1. ASVAB predicting multitasking, $\chi^2(62) = 148.858$, $p < .001$, CFI = .902, RMSEA = .088.

not account for significant variance in the context of our other predictors, and the path to attention control remained robust.

Discussion and Conclusions

In this study, we were interested in exploring ways in which we could improve the predictive validity of the ASVAB when multi-tasking was used as a proxy for real-world job performance. Moreover, we wanted to identify which types of measures would provide the best prediction. We supported the following hypotheses using a series of sequential regression analyses and structural equation models:

Fluid abilities (attention control and fluid intelligence) will account for variance in multitasking beyond that accounted for by the ASVAB (H1).

Non-subtractive accuracy-based attention tasks will have superior prediction to reaction-time difference score-based measures (H2).

Attention control will account for unique variance in multitasking above and beyond fluid intelligence (H3).

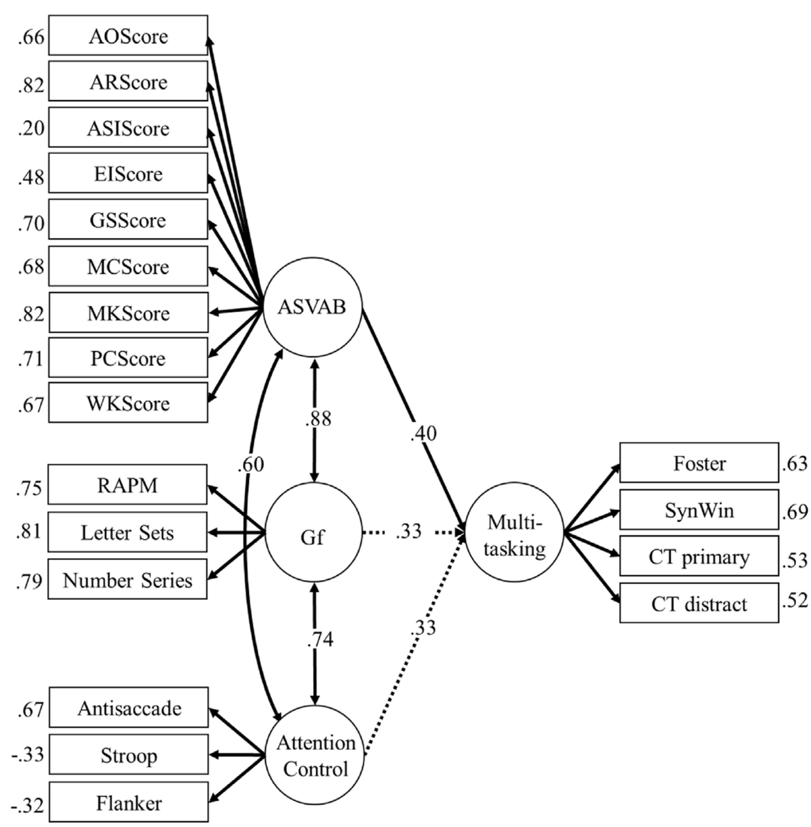

Figure 2. Latent factors of the ASVAB, fluid intelligence (Gf), and attention control were included as predictors of multitasking performance. The attention control factor was comprised of the anti-saccade and two reaction time difference score measures the Stroop and flanker. In this model, both the ASVAB and our attention control factor significantly predicted multitasking performance, $\chi^2(146) = 266.656$, $p < .001$, CFI = .909, RMSEA = .083.

Please cite this article in press as: Martin, J., et al. Improving the Validity of the Armed Service Vocational Aptitude Battery with Measures of Attention Control. *Journal of Applied Research in Memory and Cognition* (2020), https://doi.org/10.1016/j.jarmac.2020.04.002

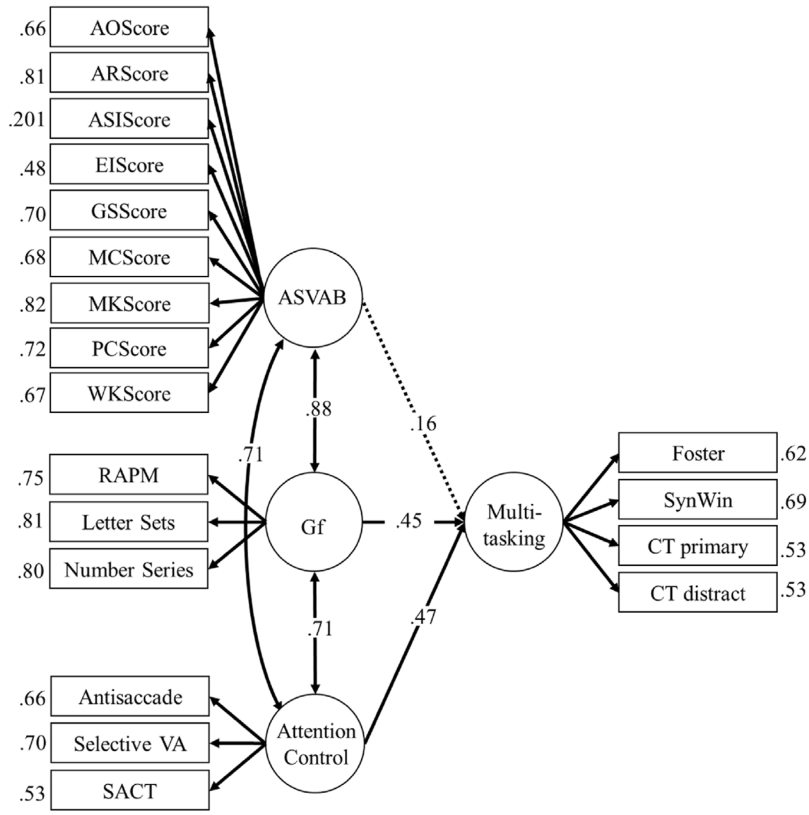

Figure 3. Correlational model with the ASVAB, fluid intelligence, attention control, and multitasking. Only the path from attention control to multitasking was significant, $\chi^2(146) = 266.027$, $p < .001$, CFI = .915, RMSEA = .07.

Processing speed will predict unique variance in multitasking above and beyond attention control if it is independently informative (H4).

In support of H1, both our regression and latent variable analyses showed that measures of Gf and attention control had incremental validity over the ASVAB. This pattern was particularly strong for the non-subtractive accuracy based tasks (antisaccade, selective visual arrays, and SACT), lending support to H2. Moreover, our measures of attention control accounted for variance in multitasking performance above and beyond measures of Gf, suggesting that attention control is not redundant with Gf (H3). Finally, the degree to which attention control predicted unique variance in multitasking performance could not be explained in terms of speed of processing, at least as defined here (H4).

Our results are important for several reasons. First, we replicated the existence of a coherent multitasking latent factor, first reported by Redick et al. (2016). Additionally, we showed a strong relationship between the ASVAB and multitasking ability at the latent level, not merely the task level (Figure 1). Once the amount of variance the ASVAB predicted was allowed to vary with Gf and attention control, as measured by accuracy-based tasks, only fluid abilities (attention control and Gf) significantly predicted multitasking ability (Figure 3). This lack of a significant relationship between the ASVAB and multitasking when accuracy-based attention control measures were used implies that the proportion of variance in multitasking predicted by the ASVAB was largely based on fluid processes related to the ability to control attention. This conclusion was not surprising given what we know about attention control and its relationship to higher cognition (Shipstead et al., 2016).

However, we also showed that the degree to which attention control provides unique prediction beyond the ASVAB depends on the way in which attention control is measured. In the regression analyses, reaction time difference score measures added no predictive validity. Similarly, a latent attention control variable defined in part by reaction time difference score tasks did not uniquely predict multitasking at the latent level (Figure 11). When the non-difference score alternatives were included, the story changed substantially, and attention control gained predictive power. We suspected that prior failures to find a full mediation of the relationship between Gf and multitasking ability (cf. Redick et al., 2016) could have been related to this difference-score issue, but this was not the case. Both attention control and Gf, while related, independently added prediction to multitasking beyond that provided by the ASVAB.

We also addressed a potential criticism that our non-subtractive accuracy based attention control tasks merely reflect individual differences in processing speed (cf. Rey-Mermet et al., 2019). If individual differences in processing speed were responsible for performance on the accuracy-based measures of attention control, then the predictive path from the processing speed latent factor should have superseded the one from attention control to multitasking. Rather, the opposite occurred: when processing speed was added to the model, it did not provide any significant predictive value above and beyond attention control. These results suggest that the ability to control and manipulate attention, and not merely the ability to process stimuli more quickly, is the fundamental component of prediction for our attention control measures.

These results support the use of basic cognitive measures, and attention control in particular, in selection contexts. Previous work from our lab on fluid abilities such as working memory capacity and Gf have been adopted by the private sector[6] and other areas of the United States Military.[7] The contribution of the present study in particular is in the simple nature of the measures of attention control and the resulting ease of understanding the individual task demands, compared to more complex problem solving tasks (e.g., Gf). In military contexts, the intended target of this proof of concept study, the ability to improve selection is particularly important given the amount of time and money that goes into training individuals for active duty and other vocational positions. As such, the ability to administer these measures quickly, easily, and efficiently, while also improving selection and retention could result in significant savings, both in terms of time and taxpayer money. As we discussed in the introduction, not only are we interested in selecting individuals who are capable of currently performing various jobs within the military, but also individuals who will be able to adjust to changing workplace demands, scope, and technological advances. The addition of fluid measures will aid in this latter aspect of selection, which will not only improve personnel selection but ideally reduce attrition as well resulting in a long range impact on the resources necessary to train and supervise future recruits.

Author Contributions

Dr. Randall Engle conceived of the study design and was the PI on the grant used to fund the project. Jessie Martin and Cody Mashburn oversaw data collection and analysis. Dr. Martin was primarily involved in the SEM analyses and Mr. Mashburn was largely involved in the regression and adverse impact analyses in the appendix. Dr. Martin wrote the majority of the paper (introduction, discussion, and results) with input and analyses from Mr. Mashburn. Dr. Engle reviewed all relevant writing and analyses prior to submission.

[6] See link to AON gamified tasks created in collaboration with Dr. Randall Engle: https://assessment.aon.com/enus/assessment-solutions/gamified-assessment

[7] Work from Dr. Engle's lab has been used to add measures of fluid abilities to the Defense Language Aptitude Battery (DLab) in a revision of their selection materials (DLab2).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jarmac.2020.04.002.

References

Barron, L. G., & Rose, M. R. (2017). Multitasking as a predictor of pilot performance: Validity beyond serial single-task assessments. *Military Psychology*, *29*, 316–326. http://dx.doi.org/10.1037/mil0000168

Bollen, K. A., & Long, J. S. (Eds.). (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. Amseterdam, Netherlands: Elsevier Science Publishers B.V.

Cepeda, N. J., Blackwell, K. A., & Munakata, Y. (2013). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, *16*, 269–286. http://dx.doi.org/10.1111/desc.12024

Colom, R., Martínez-Molina, A., Shih, P. C., & Santacreu, J. (2010). Intelligence, working memory, and multitasking performance. *Intelligence*, *38*, 543–552. http://dx.doi.org/10.1016/j.intell.2010.08.002

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100. http://dx.doi.org/10.1016/j.cogpsych.2004.12.001

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684. http://dx.doi.org/10.1037/h0043943

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review of problems and alternatives. *Psychological Bulletin*, *45*, 508–535. http://dx.doi.org/10.1037/bul0000192

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R.W. (in press) A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Elsmore, T. F. (1994). SYNWORK: APC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instrumentation, and Computers*, *26*, 421–426. http://dx.doi.org/10.3758/BF03204659

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. http://dx.doi.org/10.1111/1467-8721.00160

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*, 190–193. http://dx.doi.org/10.1177/1745691617720478

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

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*, 373–378. http://dx.doi.org/10.1111/j.0956-7976.2004.00687.x

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. http://dx.doi.org/10.1016/j.cortex.2016.04.023

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237–241. http://dx.doi.org/10.1111/j.1467-9280.1996.tb00366.x

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*, 1279–1296.

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, *24*, 1149–1167. http://dx.doi.org/10.1002/acp.1624

Hambrick, D. Z., Rench, T. A., Poposki, E. M., Darowski, E. S., Roland, D., Bearden, R. M., ... & Brou, R. (2011). The relationship between the ASVAB and multitasking in navy sailors: A process-specific approach. *Military Psychology*, *23*, 365–380. http://dx.doi.org/10.1080/08995605.2011.589323

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 645.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189.

Kleider-Offut, H. M., Clevinger, A. M., & Bond, A. D. (2016). Working memory and cognitive load in the legal system: Influences on police shooting decisions, interrogation, and jury decisions. *Journal of Applied Research in Memory and Cognition*, *5*, 426–433. http://dx.doi.org/10.1016/j.jarmac.2016.04.008

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.

Kvist, A. V., & Gustafsson, J. E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence*, *36*, 422–436. http://dx.doi.org/10.1016/j.intell.2007.08.004

Martin, J. D., Tsukahara, J. S., Mashburn, C. A., Draheim, C., Shipstead, Z., Vogel, E. K., & Engle, R. W. (under revision). *The visual arrays task: Visual storage capacity or attention control?* Manuscript submitted for publication.

Matthews, G., Warm, J. S., Shaw, T. S., & Finomore, V. S. (2014). Predicting battlefield vigilance: A multivariate approach to assessment of attentional resources. *Ergonomics*, *57*, 856–875. http://dx.doi.org/10.1080/00140139.2014.899630

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*, 222–243. http://dx.doi.org/10.1037/a0017619

Morrow, D. G., Menard, W. E., Ridolfo, H. E., Stine-Morrow, E. A. L., Teller, T., & Bryant, D. (2003). Expertise, cognitive ability, and age effects on pilot communication. *The International Journal of Aviation Psychology*, *13*, 345–371. http://dx.doi.org/10.1207/S15327108IJAP1304_02

National Research Council. (1999). *The changing nature of work: Implications for occupational analysis*. Washington, DC: The National Academies Press. http://dx.doi.org/10.17226/9600

Powers, R. (2011). *ASVAB for dummies* (3rd ed.). Indianapolis, IN: Wiley Publishing, Inc.

Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.

Redick, T. S. (2016). On the relation of working memory and multitasking: Memory span and synthetic work performance. *Journal of Applied Research in Memory and Cognition*, *5*, 401–409. http://dx.doi.org/10.1016/j.jarmac.2016.05.003

Redick, T. S., Shipstead, Z., Meier, M. E., Montroy, J. J., Hicks, K. L., Unsworth, N., ... & Engle, R. W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and fluid intelligence. *Journal of Experimental Psychology: General*, *145*, 1473–1492. http://dx.doi.org/10.1037/xge0000219

Redick, T. S., Unsworth, N., Kelly, A. J., & Engle, R. W. (2012). Faster, smarter? Working memory capacity and perceptual speed in relation to fluid intelligence. *Journal of Cognitive Psychology*, *24*, 844–854. http://dx.doi.org/10.1080/20445911.2012.704359

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of the ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement*, *54*, 459–463. http://dx.doi.org/10.1177/0013164494054002020

Rey-Mermet, A., Gade, M., Souza, A. S., Bastian, C. C., & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*, http://dx.doi.org/10.1037/xge0000593

Roberts, R. D., Goff, G. N., Anjoul, F., Kyllonen, P. C., Pallier, G., & Stankov, L. (2000). The Armed Services Vocational Aptitude Battery (ASVAB): Little more than acculturated learning (Gc)!? *Learning and Individual Differences*, *12*, 81–103. http://dx.doi.org/10.1016/S1041-6080(00)00035-2

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*, 374–393. http://dx.doi.org/10.1037/0096-3445.123.4.374

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763–776. http://dx.doi.org/10.1037/0012-1649.27.5.763

Shaw, T. H., Matthews, G., Warm, J. S., Finomore, V. S., Silverman, L., & Costa, P. T. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research in Personality*, *44*, 297–308. http://dx.doi.org/10.1016/j.jrp.2010.02.007

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*, 771–799. http://dx.doi.org/10.1177/1745691616650647

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116–141. http://dx.doi.org/10.1016/j.jml.2014.01.004

Shute, V. J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research*, *7*, 1–24. http://dx.doi.org/10.2190/VQJD-T1YD-5WVB-RYPJ

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

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.

Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence*, *7*, 53–70. http://dx.doi.org/10.1016/0160-2896(83)90006-5

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*, 500.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (3rd ed.)*. San Antonio, TX: Psychological Corporation.

Wood, G., Hartley, G., Furley, P. A., & Wilson, M. R. (2016). Working memory capacity, visual attention and hazard perception in driving. *Journal of Applied Research in Memory and Cognition*, *5*, 454–462. http://dx.doi.org/10.1016/j.jarmac.2016.04.009

Received 17 November 2019;
received in revised form 5 April 2020;
accepted 9 April 2020
Available online xxx