Research Article

A New Look at the Relationship Between Individual Differences in Cognitive Abilities and Automation-Assisted Performance

Human Factors
2026, Vol. 0(0) 1–20

DOI: 10.1177/00187208261485051
journals.sagepub.com/home/hfs
Sage

Claire Textor[1], Richard Pak[1], Anne Collins McLaughlin[2], and Randall Engle[3]

AQ1

Abstract

Objective: This study explored whether and which cognitive abilities are supported by various levels of automation.

Background: Parasuraman's (2000) model of automation provided guidance for designing automation to augment cognitive tasks. Subsequent mixed empirical findings and evolving understanding of attention control in complex tasks suggested an evaluation of how automation supports cognition within a task was needed.

Method: Participants completed cognitive ability tests then a targeting task either manually or with information automation or decision automation.

Results: Decision automation supported working memory only when reliable, while attention control predicted performance regardless of automation level.

Conclusion: Results suggest automation can support specific processes like working memory, while attention control demands may persist across automation levels. However, because the task involved a relatively constrained form of automation, future research should examine whether these effects hold with more flexible, adaptive systems.

Application: Automation should target specific cognitive functions while recognizing that attention control demands may persist. Future research might investigate how other automation designs can better support this foundational ability.

Keywords

levels of automation, attention control, working memory, individual differences

[1]Department of Psychology, Clemson University, Clemson, SC, USA
[2]Department of Psychology, North Carolina State University, Raleigh, NC, USA
[3]School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA

AQ2

Corresponding Author:
Claire Textor, Department of Psychology, Clemson University, 105 Sikes Hall, Clemson, SC 29634-0001, USA.
Email: claire.noel.textor@gmail.com

AQ3

Introduction

The primary purpose of automation is to alleviate operator workload and improve task performance, though this relationship is highly context-dependent, varying based on task characteristics, automation design, and operational environment (Jamieson & Skraaning, 2018). It may also depend on individual differences in cognition (e.g., Pak et al., 2017; Pak et al., 2024; Parasuraman et al., 2000). In the current study, we include the measurement of individual differences with an experimental approach to examine how individual variation in cognitive abilities might relate to automation-assisted performance.

When physical tasks are automated, it is relatively clear what work the automation supports. For example, a mechanical tractor replaces the movement of a hand-held scythe. However, when automation supports cognitive work (i.e., tasks with attention, memory, or decision-making components) it becomes less clear which cognitive process the automation supports. Indeed, in many cases automation claims to aid a cognitive process that it may not support (McLaughlin & Byrne, 2020). Parasuraman et al. (2000) addressed this uncertainty by categorizing automation into four types distinguished by the cognitive processes they support, ranging from information acquisition (perception and attention) through information analysis (working memory [WM]) and decision support (decision making and response selection) to action implementation.

Within each type, a task can be more or less automated, referring to its level of automation (Parasuraman et al., 2000). The combination of type and level is the *degree of automation,* where higher degree automation assumes more of a task that is cognitive in nature, for example, by offering potential decisions to the operator (Onnasch et al., 2014). Parasuraman et al.'s (2000) framework has provided researchers and practitioners with a more human-centered model for understanding and designing automation from a psychological perspective, rather than one centered on machine capabilities.

If the precise cognitive ability requirements of a task are known (e.g., through task analysis) and an appropriate level of automation is introduced that targets that ability according to Parasuraman et al.'s (2000) framework, individual differences in that required ability (e.g., WM) should no longer be a limiting factor to performance. For example, navigating to a new destination can be WM-intensive because the driver must keep route information temporarily accessible while using it to make ongoing driving decisions. Therefore, individuals with higher WM may outperform those with lower WM. However, the introduction of a navigation aid that provides decision automation (e.g., a global positioning system [GPS]), designed to alleviate WM demands, should raise the performance of low-WM individuals to or near the level of high-WM individuals, eliminating the relationship between WM and performance.

However, the introduction of automation can have unintended effects on human performance (Bainbridge, 1983) by altering the nature of the task in unexpected ways (Endsley & Kaber, 1999; Wright & Kaber, 2005). For example, a GPS may eliminate the need to hold information for short periods of time but could also inadvertently increase the need to switch attention between the road and device display. The automation may also alter the task by requiring the human to monitor its performance and take over control following failures. Even though high-level decision can assume some parts of a task, it does not supplant the operator's role. The operator's responsibilities are altered to be more supervisory, often requiring vigilant monitoring and intervention in case of a failure. Such changes in operator demands and a lack of focus on those changes in research designs are evidenced by mixed findings in the literature, with some research showing a positive relationship between WM and automation-assisted performance (e.g., Rovira et al., 2017) and others finding no significant relationship (Pak et al., 2017). This uncertainty

may be due to two inter-related factors: different experimental designs, and an incomplete understanding of the role of human cognitive abilities in task performance and how automation supports those abilities. While task- and system-specific details also affect performance with automation (e.g., Skraaning & Jamieson, 2021), the present investigation focused on the role of individual differences in cognitive abilities on automation-assisted performance.

Working Memory and Automation Performance

Working memory has been one of the most studied cognitive abilities in human factors (Pak et al., 2024) and also in automation research with an individual differences focus. Research has demonstrated a positive relationship between WM and performance with automation (de Visser et al., 2010; Rovira et al., 2017). For example, de Visser et al. (2010) found that WM predicted performance in an automated air defense task even with automation present, suggesting the low-level automation (alerting) did not fully support WM demands. However, because that study only included alerting, participants were still responsible for integrating information, comparing alternatives, and choosing actions, leaving open whether higher-level automation would better support WM.

To address this limitation in the literature, Rovira et al. (2017) examined the relationship between WM and various degrees of automation (manipulating *what* and *how much* is being automated). In their task, participants were presented with a terrain display showing up to six enemy units, up to six friendly units, three battalion units, and one headquarters unit. The participant's task was to identify the best enemy-friendly pair based on proximity to each other and headquarters distance. They received either full distance calculations between all units (information automation) or a filtered list of the top three recommended pairings (decision automation), with 80% automation reliability. When information automation was unreliable, the distance calculations were inaccurate. Inaccurate decision automation suggested pairings which did not represent the true top three pairing options. Findings showed that when automation was reliable, there were no differences in accuracy between those with higher or lower WM ability. However, when the automation was unreliable (i.e., inaccurate), those with higher WM outperformed those with lower WM. This finding demonstrated that high-level automation could support the WM demands of a task, but only when it was error-free. Using the same task as Rovira et al. (2017), Pak et al. (2017) found, in an older adult sample, that level of automation was unrelated to individual differences in WM. Some potential reasons for this discrepancy may be that each study used different age groups (younger in Rovira et al., versus older in Pak et al.), and that each study measured WM differently (spatial memory or complex span).

These mixed findings may reflect two methodological issues. First, most studies rely on a single indicator of WM (Conway et al., 2005; see Pak et al., 2024 for a review), which is vulnerable to method variance and makes cross-study comparison difficult given the variety of measures used (cf., Operation Span [OSPAN] (Turner & Engle, 1989) and Spatial WM (Corsi, 1972). Second, introducing higher levels of automation may alter the cognitive requirements of the task in unanticipated ways, potentially shifting demands from memory to attentional processes such as monitoring (Endsley & Kaber, 1999).

In addition to the above methodological issues present in the existing research, there have been recent theoretical findings on the nature of WM and its role in complex task performance. Earlier, it was assumed that WM and reasoning ability, or general fluid intelligence (Gf) were essentially equivalent, with both being highly predictive of complex task performance (Kyllonen & Christal, 1990). However, more recent findings have shown that WM, Gf, and by extension complex task performance, may all be influenced by a

higher-level cognitive construct: attention control (Burgoyne et al., 2023; Draheim et al., 2022; Pak et al., 2024).

Attention Control and Automation-Assisted Performance

Attention control (AC) is a fundamental cognitive capability that shapes how effectively individuals can regulate their information processing to achieve goals (Burgoyne & Engle, 2020; Burgoyne et al., 2023; Oberauer, 2024). AC is thought to exert its effects on the flow of goal-directed thought by controlling two critical functions: WM that maintains goal-relevant information in an active state and Gf that resolves goal conflict and discards irrelevant information (Shipstead et al., 2016). Research has revealed that AC orchestrates the deployment of WM and Gf (Burgoyne & Engle, 2020; Burgoyne et al., 2023), and accounts for the close statistical relationship between them (i.e., WM and Gf are no longer correlated when variance in AC is taken into account; Draheim et al., 2021). Because of this central role, AC serves as the primary predictor of multitasking performance, explaining approximately 75% of variance in complex multitasking ability (Burgoyne et al., 2023). This hierarchical position of AC above other cognitive abilities, combined with its close relationship to WM and dominant influence on multitasking performance, raises a critical question: does automation designed to support WM also support AC abilities?

Given that automation is often deployed in complex, multitasking situations, there is reason to believe that AC plays an important role in automation-assisted task performance. It is plausible that the functions controlled by AC (maintenance of goal-relevant information and disengagement from irrelevant information) should play a major role in how individuals perform with automation. More specifically, automation monitoring requires sustained attention to detect failures or deviations from expected outcomes. Those who are better able to ignore irrelevant information (Gf-dependent) and maintain task-relevant details (WM-dependent) should be better able to detect failures if they have a high capacity to flexibly switch between those two functions (AC-dependent). Those who find this task challenging may struggle to detect failures, leading to an inaccurate perception of performance and subsequent over trust in the automation. In the event of a failure, individuals must be ready to disengage from the automation's output and resume the task manually, a process that draws directly on AC's dual functions of disengagement and goal maintenance.

Overview of the Current Study

Parasuraman et al.'s (2000) model for the types of automation provided an initial framework for determining which system functions should be automated and to what extent, with the resulting consequences for human performance serving as an important basis for evaluating automation design (Pak et al., 2017; Parasuraman et al., 2012, 2014; Rovira et al., 2007, 2017). However, methodological limitations (i.e., disparate use of WM measures, using single indicators instead of latent factors) may be partially responsible for disparate findings. Using the same automation-assisted decision task as prior studies (Pak et al., 2017; Rovira et al., 2007, 2017), the current study addresses some of the methodological limitations surrounding the measurement of WM by using two complex span tasks that better measure its inherent storage and processing nature (La Pointe & Engle, 1990). AQ4

As part of our broader investigation, we also explored how variations in Gf and AC may relate to automation-assisted performance. While prior research had shown a strong relationship between WM and Gf (e.g., Kyllonen & Christal, 1990), more recent research (Draheim et al., 2021) has shown that this relationship is fully mediated by AC. However, current automation research examining individual differences largely

centered on WM (e.g., Rovira et al., 2017; Figure 1(B)). This narrow focus on WM overlooked potential interactions with other related abilities. Given the established interplay between AC, WM, and Gf (i.e., Draheim et al., 2021), a more comprehensive approach necessitates exploring how these additional cognitive abilities collectively impact automation-assisted performance.

In general, consistent with Parasuraman et al.'s (2000) model, we expected that individual differences in WM would only be related to performance with unreliable automation (Rovira et al., 2017). Conversely, we expected that reliable high-level automation would support working memory demands for those with lower capabilities, thus eliminating any significant relationship between working memory and performance. However, given the importance of AC for complex task performance, our second exploratory goal was to examine the extent to which existing automation accommodated AC demands. In the context of this paper, "existing automation" refers to conventional systems that support predefined stages of task performance, for example by presenting processed information or decision recommendations. Given AC's role in modulating both WM and Gf and its demonstrated importance in multitasking (Burgoyne et al., 2023), we tentatively expected that AC would predict performance and that this relationship might not be fully accommodated by existing automation designs developed primarily to support WM.

Method

Design

The study employed a between-participants design with three experimental conditions varying in automation level: manual control (no automation), Information Automation (IA), and Decision Automation (DA). In both automation conditions (IA and DA), system reliability was set at 80% after the first 10 trials, meaning that automated suggestions were accurate in 80% of trials and inaccurate in 20% of trials. Participants were randomly assigned to one of the three conditions. Working Memory, Gf, and AC were treated as continuous predictor variables. Each factor was computed as a composite score by averaging standardized (z-score) performance across the multiple cognitive ability tests measuring that construct. Working memory was derived from two complex span tests (short OSPAN, symmetry span; Foster et al., 2015; Unsworth et al., 2009), while AC combined three attention control tests (Burgoyne et al., 2023). Fluid intelligence

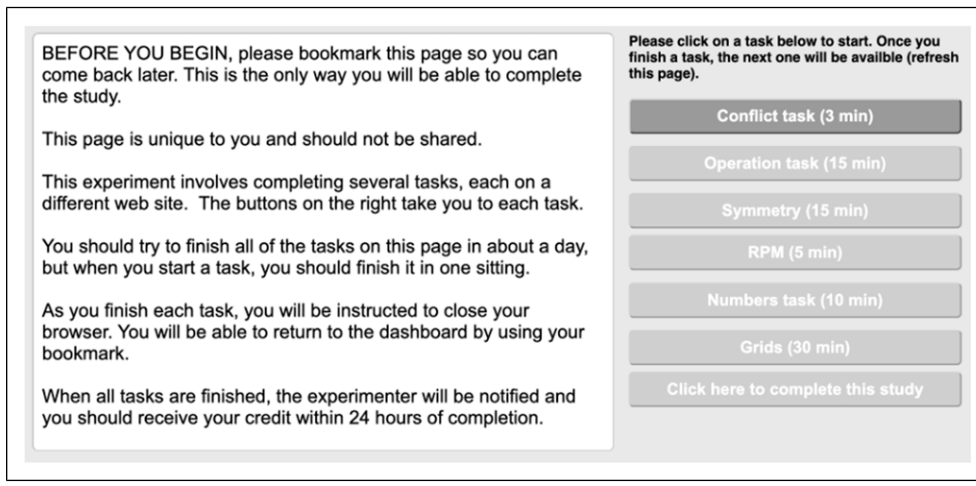

Figure 1. Experimental dashboard (with task duration estimates)

combined performance on Raven's Progressive Matrices (Kane et al., 2004; Raven & Court, 1998) and Number Series (Thurstone, 1938; Unsworth et al., 2009) tests.

Participants

An *a priori* power analysis was conducted assuming 80% power and a small-to-moderate effect size (Cohen's $f = 0.20$). Under these assumptions, the required sample size was 102 participants (Faul et al., 2007). A convenience sample of 318 college students ranging in age from 18 to 32 were recruited and received partial course credit. Because the study was administered online, data were screened for inattentive responders and technical errors. Ninety-one participants were excluded from analysis because they scored below chance or provided no data on at least one cognitive ability test (described below), a rate consistent with previous studies using these online measures (Burgoyne et al., 2023, 2024; Draheim et al., 2021; Mashburn et al., 2024). Twenty-four additional participants were excluded due to performance below 10% on the experimental task. The final sample consisted of 203 individuals (151 females, $M_{age} = 18.9$, $SD = 2.6$). This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Clemson University. Informed consent was obtained from each participant.

Task and Procedure

After providing consent, participants accessed a web dashboard with links to each of the cognitive ability tests and the experimental task (Figure 1). Since the dashboard link was unique and persistent to each participant, they could complete the study at their own pace (e.g., finish one task, take a break, and begin the next task). Completing all tests and the experimental task took about 60 to 80 min.

Cognitive Ability Tests. Participants completed a battery of cognitive ability tests: Three conflict squared tasks (shown as "conflict task" in Figure 1), which are Stroop-like tests of attention control (Burgoyne et al., 2023). We measured working memory using two complex span tests: shortened OSPAN ("operation task"), which required participants to remember letters while solving math equations, and symmetry span ("symmetry"), where participants remembered spatial locations while judging pattern symmetry (Foster et al., 2015). Finally, fluid intelligence was assessed using Raven's Progressive Matrices ("RPM"; Kane et al., 2004; Raven & Court, 1998), and the number series task ("numbers task"; Unsworth et al., 2009; Thurstone, 1938). In both of the tasks, participants predicted the next likely sequence in a series of patterns (Raven's) or numerals (number series). Each participant completed the cognitive ability tests in a fixed order and had to finish each task before moving onto the next task. Prior studies have found that remote versions of these measures produce psychometric properties and score distributions comparable to in-person administration (Burgoyne et al., 2023, 2024, in press; Mashburn et al., 2024).

Experimental Task. The experimental task ("Grids", Figure 2) was a web-based, low-fidelity military unit dispatching simulation used in prior automation research (Pak et al., 2017; Parasuraman et al., 2012, 2014; Rovira et al., 2007, 2017). The interface consisted of two sections: an overhead terrain grid displayed on the right side showing all elements, and a targeting input area positioned at the top left. For conditions where automation was present, automated assistance was displayed in the middle-left section (Figure 2(A)–(C)).

The terrain grid contained 16 elements across all 100 trials: 6 enemy blocks (colored red), 6 artillery blocks (colored green), 3 battalion blocks (colored yellow serving only as visual distractors), and 1 headquarters (HQ) block (colored orange), all positioned randomly for each trial. Participants identified the enemy-artillery pair that satisfied two criteria: (1) closest to each other and (2)

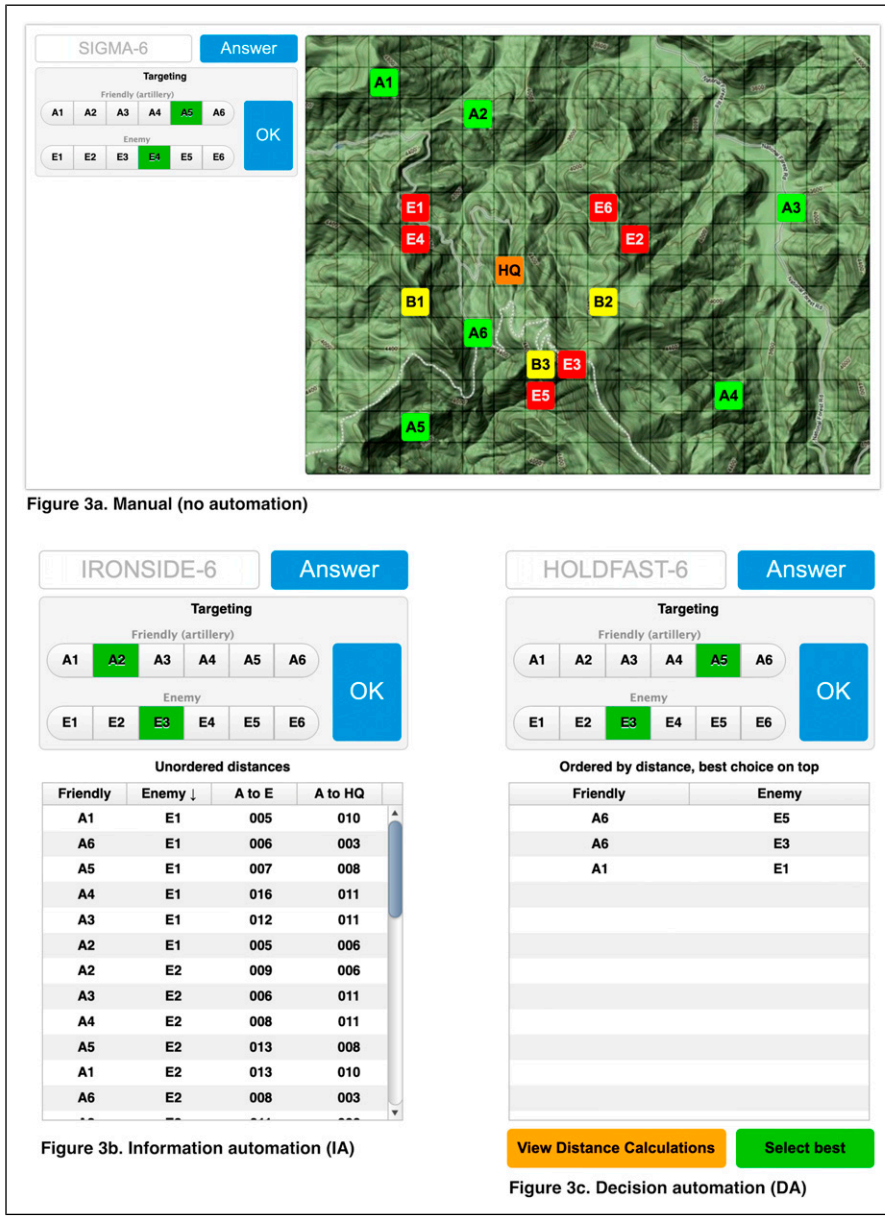

Figure 3a. Manual (no automation)

Figure 3b. Information automation (IA)

Figure 3c. Decision automation (DA)

Figure 2.(A)–(C). The three between-participant conditions (all conditions saw the grid-like map shown in 3(A)). Figure 2(A) shows the manual condition, with a targeting panel for choosing the target-enemy pair. Figure 3(B) shows the IA condition (cropped to remove terrain grid which was identical) where all distances between possible squares are calculated but unordered. Figure 2(C) (cropped) shows the DA condition which showed only 3 target-enemy pairs with the best one (using the criteria of distance and HQ proximity) at the top. Numerical distances (A–E and A–HQ) were hidden to reduce distraction, but participants could view them ("View distance…" button) or simply choose the top recommendation ("Select best" button), or manually input their target pair

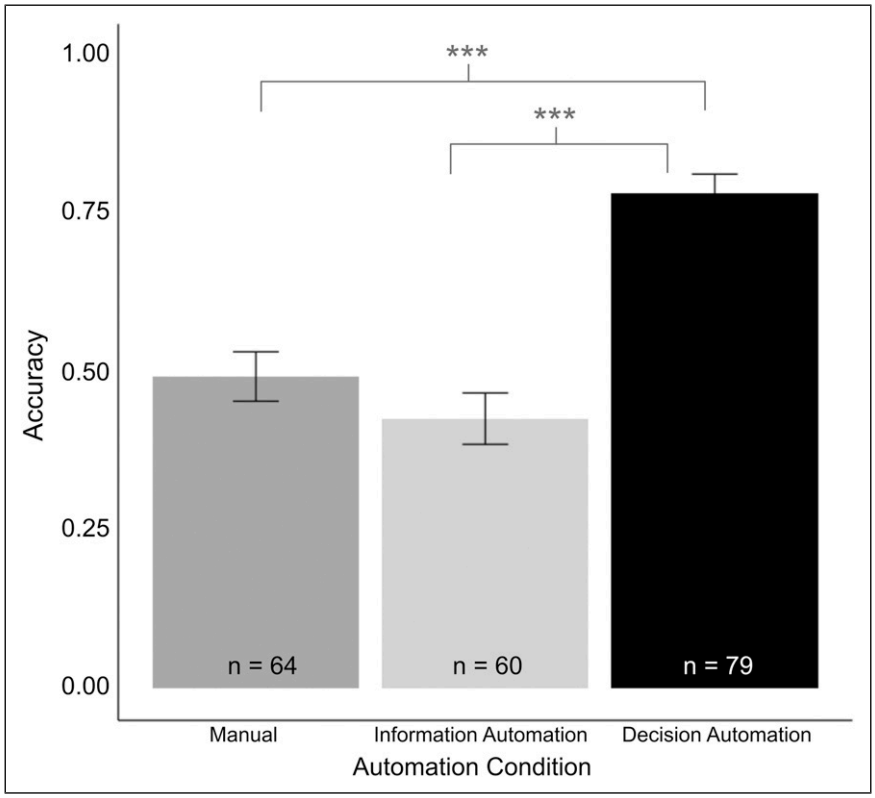

Figure 3. Mean accuracy in the three between-participant automation conditions with number of participants indicated for each. Bars represent 95% confidence intervals. *Post hoc* comparisons showed differences between decision automation performance and the other conditions; *$p < .05$, **$p < .01$, ***$p < .001$

nearest to HQ. When multiple pairs had equal distances, participants prioritized the enemy closest to headquarters. Selections were made using the "Targeting" panel on the left side (Figure 2(A)–(C)), with confirmation via an "OK" button to advance to the next trial. Each trial was limited to 11 s; timed-out trials advanced automatically.

Automated Aids: Information and Decision Automation. When in an automation condition, automation-assisted participants in choosing enemy-friendly pairs. The information automation (IA) condition displayed an alphabetically ordered list of all pairs with their distances from each other and from headquarters. Participants needed to scroll through this list and apply selection criteria to find the optimal pairing (Figure 2(B)). During unreliable trials, inaccurate distances were displayed.

The Decision Automation (DA) displayed only the top three recommended pairs in ranked order, without numerical distance values (Figure 2(C)). When reliable, it correctly calculated distances and applied the selection criteria (closest A–E pair that was also closest to HQ). During unreliable trials, it presented random recommendations. This qualified as decision automation because it substituted machine recommendations for human judgment. Participants could verify the recommendations by revealing the underlying distance calculations and selecting a pair manually or comply with the automation by accepting the top-ranked recommendation.

Both IA and DA operated at 80% reliability. This value was consistent with prior applications of the task (Pak et al., 2017; Rovira et al., 2007, 2017) and falls above the approximately 70% threshold at which automation benefits outweigh costs (Wickens & Dixon, 2007), while providing sufficient unreliable trials to examine individual differences in failure detection. No failures occurred during the first 10 trials, allowing participants to establish initial trust.

Participants simultaneously monitored call signs in a Communications panel. This secondary task increased overall workload (incentivizing automation use when available) and functioned as an attention check. The communications panel cycled through 14 possible call signs every 6 s (with replacement), requiring participants to click an Answer button specifically when "MASON-6" appeared (probability 1/14 per cycle). Participants were instructed to allocate equal attention to both the communications and experimental tasks. Performance on this secondary task was high across all experimental conditions (see Table 1 under Supplemental Materials). DA verification and compliance rates are reported in Table 2 under Supplemental Materials.

Prior to the experimental trials, participants completed six practice trials in their assigned condition. In conditions featuring automation, participants were informed that the computer assistance was generally reliable but not perfect, although during practice the automation maintained 100% reliability to establish initial trust.

Results

Data Preparation

Prior to analysis, accuracy data were aggregated, creating a mean for a participant's reliable trials (composed of approximately 80 trials) and one for unreliable trials (approximately 20 trials). For analyses, each mean was weighted by the number of valid trials contributing to it, thereby accounting for the unequal numbers of reliable and unreliable trials.

Due to the between-participant manipulation of Automation Condition, the resulting accuracy data did not meet the assumption for normality as assessed by the Shapiro–Wilk test ($W = 0.95$, $p < 0.001$). The models also showed evidence of heteroskedasticity. The source was that those offered DA were more accurate and less variable than the manual or IA groups. Transforming the dependent measure of accuracy with a log transform did not result in meeting the assumptions. Thus, all models (e.g., using lmer in R) were also run using the *robustlmm* package (rlmer) (Koller & Stahel, 2022). In every case, significant and non-significant findings remained the same using rlmer, and thus the most simple analyses using lmer are presented and assumed to have been robust to the violations of assumptions.

Description of the Analyses

Given that only the automation conditions were subject to the manipulation of trial-level automation reliability, a full factorial design was not possible. Consequently, our analyses focused on subsets of the complete dataset. The initial analysis compared differences across the between-participants independent variable, Automation Condition, which included three groups: Manual, IA, and DA. The variable Trial Reliability was excluded for this analysis, as there was no reliability manipulation in the manual condition.

The subsequent analysis encompassed only the IA and DA conditions and incorporated the within-participants variable of Trial Reliability. All participants in the automation conditions experienced 80% reliability overall. However, rather than treating reliability as a system-level variable, we followed Rovira et al. (2017) in categorizing each trial as either reliable (automation correct) or unreliable (automation incorrect). This is because even though all participants experienced the same overall automation reliability of 80%, reliability was analyzed at the trial level by classifying each trial according to whether the

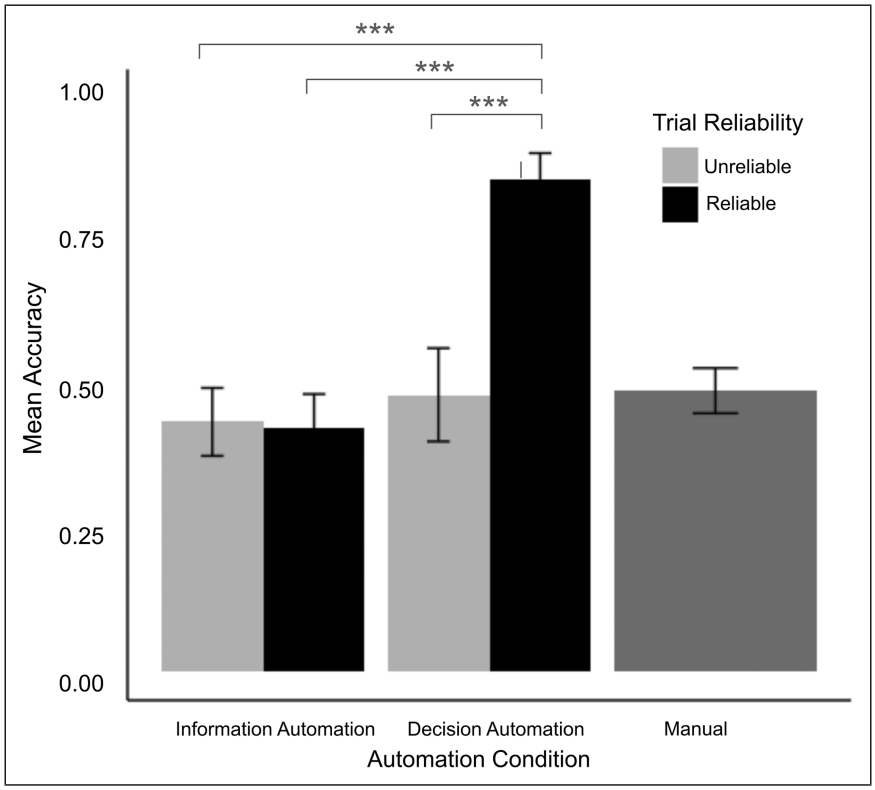

Figure 4. Mean accuracy by automation condition, with information and decision automation divided into trial reliability levels. Bars represent 95% confidence intervals. *$p < .05$, **$p < .01$, ***$p < .001$. Manual group is shown for comparison but was not included in analyses involving trial reliability

automation was correct or incorrect. This allowed us to examine whether cognitive abilities predicted accuracy differently on reliable versus unreliable trials within the same participant.

Regarding trial completion, some trials timed out, leading to a variation in the number of trials completed by participants (ranging from 80 to 100 trials). However, these time-outs were distributed evenly across the three groups (IA: 4.7% time-outs, DA: 4%, Manual: 4.6%) and across trial reliabilities (Reliable: 5.7%, Unreliable: 4.2%). This uniform distribution implied that the accuracy measurements for each group and condition were comparably influenced by the inclusion or exclusion of timed-out trials. Therefore, we decided to exclude timed-out trials from our analyses, allowing participant accuracy to reflect their performance on completed trials only. The results are presented by dependent measure.

Accuracy Across Automation Conditions

A one-way ANOVA was conducted to compare accuracy among the three between-participants automation groups: manual, IA, and DA. The analysis revealed significant differences in accuracy across these groups, $F(2, 200) = 38.4$, $p < .0001$, $\eta^2 = 0.277$. Subsequent pairwise comparisons indicated that accuracy (performance) in the DA condition was significantly better than the manual and IA conditions (all ps $< .0001$). However, there were no significant differences between the manual and IA conditions ($p = 0.08$; Figure 3). Having the DA significantly aided

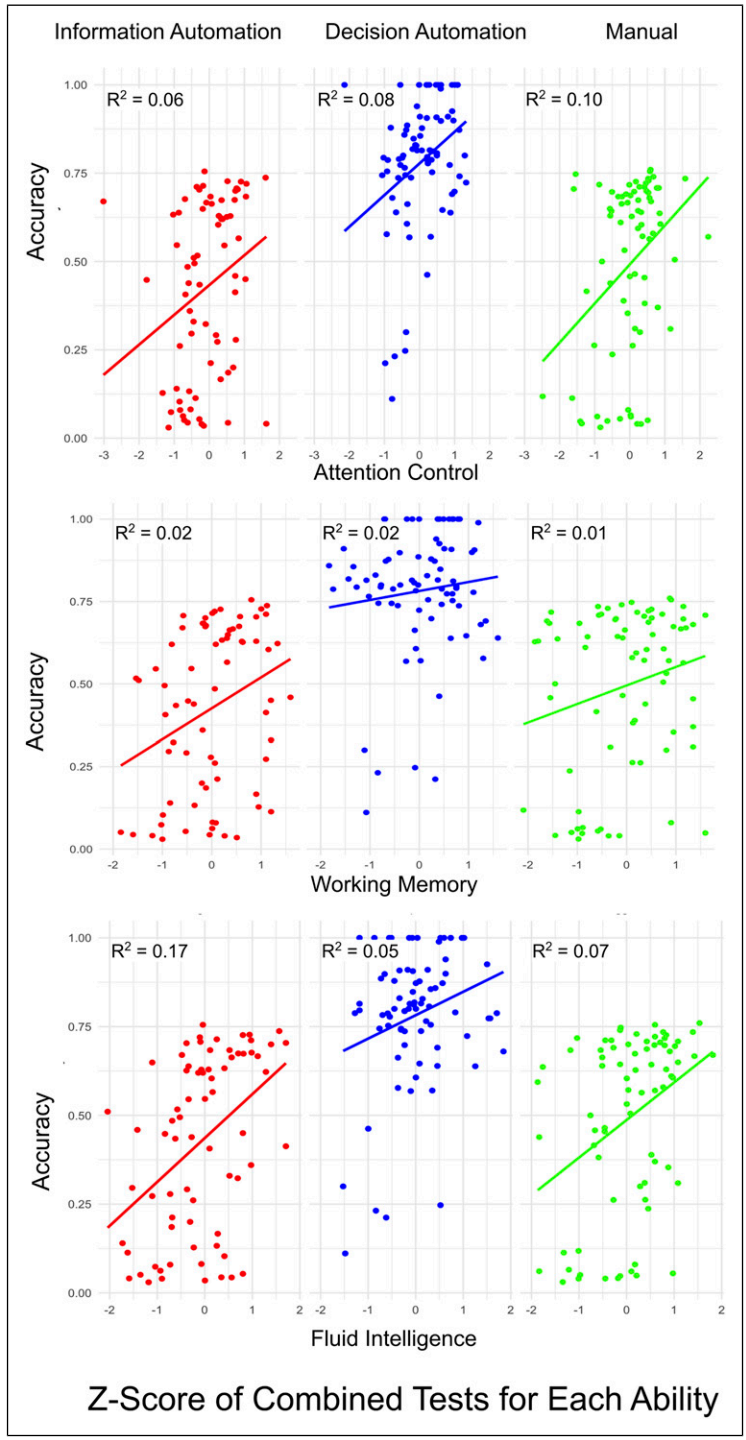

Figure 5. Correlations between the mean of combined z-scores for each cognitive ability test and accuracy for each automation condition

AQ5

Table 1. Multilevel models for predicting accuracy for the two types of automation

	Model 1: Unconditional model				Model 2: Cognitive ability test scores (Level 1)				Model 3: Manipulated variables and interactions				Model 4: Reduced to significant predictors			
Predictors	*Estimates*	*SE*	*t*	*p*	*Estimates*	*SE*	*t*	*p*	*Estimates*	*SE*	*t*	*p*	*Estimates*	*SE*	*t*	*p*
Intercept	***0.657	0.021	31.80	<0.0001	***0.655	0.020	32.99	<0.0001	***0.514	0.039	13.130	<0.001	***0.510	0.040	12.920	<0.000
Working memory (WM)					−0.009	0.027	−0.320	0.752	−0.016	0.055	−0.290	0.769	0.014	0.051	0.280	0.781
Attention control (AC)					**0.065	0.027	2.38	0.019	0.070	0.050	1.400	0.164				
Fluid intelligence (Gf)					*0.067	0.029	2.30	0.023	0.018	0.054	0.330	0.743				
Automation condition (AutoCond)									−0.044	0.052	−0.850	0.396	−0.037	0.052	−0.720	0.474
Trial reliability (TR)									−0.015	0.037	−0.390	0.695	−0.017	0.037	−0.450	0.652
AutoCond * TR									***0.390	0.049	7.900	<0.001	***0.393	0.049	8.100	<0.000
WM*AutoCond									0.121	0.072	1.700	0.091	0.115	0.067	1.710	0.089
WM*TR									0.016	0.053	0.310	0.755	0.031	0.048	0.650	0.516
Gf*AutoCond									−0.025	0.070	−0.360	0.722				
Gf*TR									0.023	0.048	0.480	0.634				
AC*AutoCond									0.038	0.075	0.500	0.615				
AC*TR									0.019	0.052	0.380	0.707				
WM*AutoCond*TR									M-0.133	0.068	−1.940	0.054	*−0.145	0.063	−2.320	0.022
Gf*AutoCond*TR									−0.025	0.067	−0.370	0.715				
AC*AutoCond*TR									−0.002	0.072	−0.030	0.980				
Random effects																
σ2	0.034				0.029				0.024				0.028			
τ00	0.129				0.129				0.065				0.064			
ICC	0.207				0.182				0.269				0.307			
N	139 participant				139 participant				139 participant				139 participant			
Observations	278				278				278				278			
Marginal R2/Conditional R2	0.999/0.207				0.999/0.182				0.999/0.269				0.999/0.307			
AIC	153.4				160.7				79.26				44.08			

Note. Marginal R2 reports variance explained by fixed factors and conditional R2 reports variance explained by both fixed and random factors. *M = marginal significance, $p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

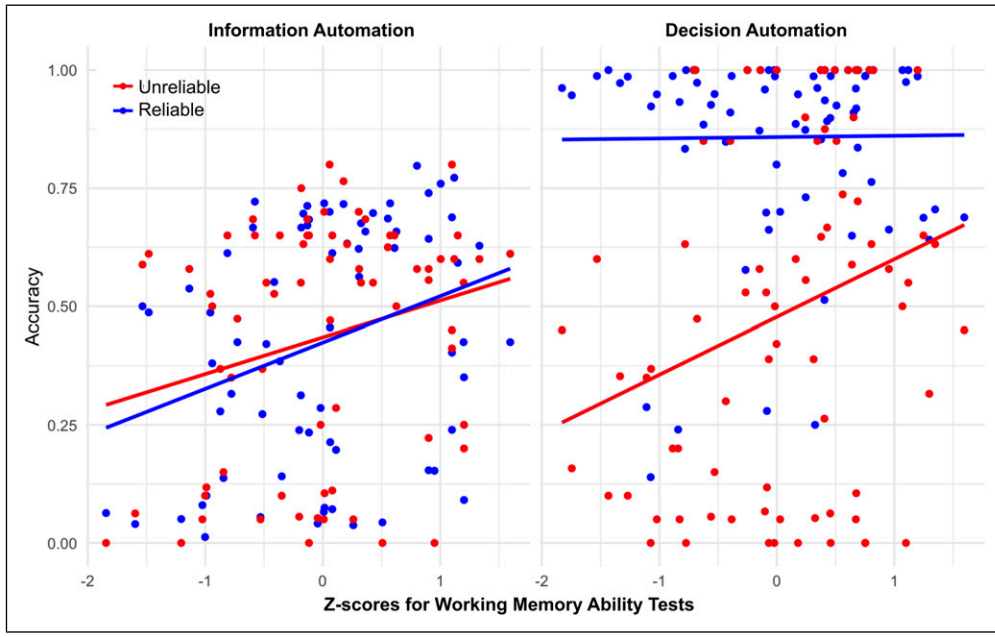

Figure 6. The three-way interaction of automation condition, trial reliability, and WM on accuracy. The *x*-axes show scores on the composite measure of WM created by averaging the standardized scores of two tests of WM

performance while the IA was no better than completing the task without automation.

Trial Reliability. We analyzed the two conditions where automation was present, divided based on Trial Reliability, categorizing trials as either reliable or unreliable. A 2x2 ANOVA was performed to investigate the impact of Automation Condition and Trial Reliability on accuracy. Given the unequal distribution of trial types (80% reliable and 20% unreliable), we applied weights to the means of each condition to balance the disparity in trial frequency. The analysis revealed significant main effects for both Automation Condition, F (1, 274) = 24.8, $p < .0001$, $\eta^2 = .065$, and Trial Reliability, $F(1, 274) = 43.2, p < .0001, \eta^2 = .114$. Furthermore, there was a significant interaction between Automation Condition and Trial Reliability, $F(1, 274) = 37.1, p < .0001$, $\eta^2 = .098$.

Post-hoc comparisons using Tukey's HSD test identified significant differences in accuracy. Participants in the DA group experiencing reliable automation exhibited significantly higher accuracy compared to all other groups (all *p*s comparing reliable-DA to other conditions <.001; Figure 4).

Individual Cognitive Differences. An initial examination revealed correlations between cognitive abilities and performance (Figure 5). To assess how these relationships varied across Automation Condition and Trial Reliability, a multilevel modeling approach, as suggested by Raudenbush and Bryk (2002), was utilized. Because Trial Reliability was manipulated within-subjects, it was nested within each participant and thus, within each ability score. A multilevel approach was taken using hierarchical linear mixed models (Raudenbush & Bryk, 2002) to account for the nested data. These models are fitted using the

lmer function from the *lme4* package in R (R Core Team, 2021; Bates et al., 2015). Results are presented as recommended in the reporting guidelines of Huang (2022). Satterthwaite's method was used to calculate degrees of freedom (Satterthwaite, 1946). Predictor variables were not centered as all had a meaningful zero. Model comparisons were performed using Akaike information criterion, Bayesian information criterion, and likelihood ratio tests, chi-square values, degrees of freedom, and *p*-values provided.

First, a null (unconditional) model with only a random intercept for individuals (denoted by ID), was run to determine the amount of variance within and between participants. Then, predictors were added to the model sequentially and the models tested for improvement. Non-significant predictors and interactions were removed, to provide a final parsimonious model that best represented the data. In all models, weights were applied to the means for reliable and unreliable trials to account for the fact that the reliable means were composed of four times as many trials as the unreliable means. Results are discussed in the text, with all statistical information available in Table 1.

Multilevel Models. The intraclass correlation coefficient of the null model was 0.21, indicating that 79% of the variance in accuracy was between participants, while 21% was within participants. These results indicated that there was sufficient variation at each level to proceed with a multilevel analysis that included between-subject variables (Automation Condition and cognitive ability tests) and the within-subject variable of Trial Reliability (Raudenbush & Bryk, 2002).

In a second model, cognitive ability test scores were entered as predictors. Attention control and Gf showed significant main effects on accuracy; those with higher scores on these tests had higher overall accuracy when controlling for the other variables. An ANOVA confirmed that this model significantly decreased fit compared to the null model, χ^2 (3) = 15, $p < .01$.

In a third model, scores on the three cognitive ability tests were included as predictors and allowed to interact with the manipulated variables of Automation Condition and Trial Reliability. The interaction of Automation Condition and Trial Reliability was significant, as was the marginal interaction ($p < 0.05$) of WM with Automation Condition and Trial Reliability. This was the only model that changed in significance of effects when resubmitted using the *robustlmm* package: the marginal effect of the WM x Automation Condition x Trial Reliability interaction became significant (t increased from −1.94 to −2.69), and the lower level interaction of WM and Automation Condition showed significance (t increased from 1.70 to 2.16). This model also demonstrated a significant improvement in fit over the previous one, as evidenced by an ANOVA, χ^2 (12) = 161, $p < .001$.

Follow up analyses showed the source of the interaction was that WM was required for the IA condition, whether or not the automation was reliable. However, for decision automation, WM only predicted performance for unreliable automation trials (Figure 6); on reliable trials participants tended to score highly no matter what their score on the WM tests (i.e., reliable decision automation eliminated the effect of individual differences in WM). Once the manipulated variables of Automation Condition and Trial Reliability were entered, the effects of AC and Gf were no longer present.

To achieve a more parsimonious model, non-significant predictors were removed, resulting in Model 4. This final model had a significantly better fit compared to Model 3, χ^2 (8) = 22.0, $p < .01$, with significant interactions of ability, condition, and reliability (Table 1).

Discussion and Conclusion

The current study reassessed which cognitive abilities are supported by different levels of automation. There is an established taxonomy that categorizes automation by the cognitive

processes it assists (Parasuraman et al., 2000). However, empirical findings are mixed on whether key abilities like WM are actually supported by high level automation. The current results showed, consistent with Parasuraman et al.'s model, that higher levels of (reliable) automation support the WM demands of the task. This was demonstrated here by showing that individual differences in WM (measured at the latent construct level with multiple indicators) predicted significant variance in performance with IA (a type of automation designed to augment perceptual abilities) but not with DA (a type of automation designed to augment WM).

We also had a second, exploratory question in examining automation's cognitive support: "what cognitive abilities remain unsupported by automation?" Given recent theoretical developments regarding the relationship between WM and AC (Engle, 2002), we investigated whether automation designed to support WM would similarly accommodate AC—a question made critical by evidence suggesting AC's primacy over WM in complex task performance (Draheim et al., 2022; Pak et al., 2024). This is particularly relevant because AC may govern operators' ability to monitor automation and adapt to dynamic task demands, which are critical in supervisory control paradigms. Our findings revealed that while individual differences in AC predicted performance, this relationship was not moderated by automation level, suggesting that existing automation may not have fully accommodated individual differences in AC in the current task context.

The findings suggest that automation designed to support WM may not inherently address individual differences in AC. Parasuraman et al.'s model maps automation to discrete stages of information processing but does not account for AC's hierarchical role in regulating cognitive resources and maintaining task priorities under high workload (cf. Wickens, 2008). Recent work positions AC as a foundational ability that governs both the maintenance of goal-relevant information and the suppression of distractors (Burgoyne et al., 2023; Shipstead et al., 2016). While decision automation alleviated WM demands, it may have introduced requirements for vigilance and failure recovery-processes that align with AC's functions of maintenance and disengagement (Shipstead et al., 2016). This is consistent with Bainbridge's (1983) "irony of automation," where systems designed to reduce workload inadvertently shift demands to underappreciated cognitive abilities. Filtered recommendations may reduce WM load but require operators to suppress automation bias (Bahner et al., 2008; Lee & See, 2004; Skitka et al., 1999) or re-engage manually during AQ6 failures. Future models could position attention control as a mediator between automation type and performance, where high attention control enables operators to adapt to automation-induced task transformations. This interpretation aligns with Multiple Resource Theory (Wickens & Dixon, 2007; Wickens et al., 2002), which implies that automation redistributes cognitive demands across modalities and resources, but AC may moderate the success of this task reallocation.

The static (always-on, non-adaptable) nature of our automation paradigm warrants consideration. AC predicted performance across conditions, suggesting it may govern residual task demands even when memory load is reduced, consistent with the shift from direct control to supervisory monitoring (Sheridan, 1992). In dynamic or adaptive automation systems where operators switch between manual and automated control (Calhoun, 2022; Miller et al., 2011; Moray et al., 2000; Parasuraman et al., 2009), AC's role may be even more pronounced.

Overall, the findings support the practical utility of Parasuraman et al.'s (2000) model: automation designed to target WM demands did so as predicted. However, they also show that automation that supports one ability (e.g., WM) may not support other closely related abilities (e.g., AC). As automation is deployed in increasingly complex task environments, understanding which cognitive demands persist despite automation support (and for

whom) may become an important design consideration (Greenwell-Barnden et al., 2026).

If AC demands do persist across automation levels, how might automation be designed to support AC? A path forward may come from the very functions that AC is thought to directly control: *maintenance* of goal-relevant information and *disengagement* from irrelevant information (Shipstead et al., 2016). Prior work has shown the performance benefits of automation that explicitly support operator's task disengagement (e.g., Dehais et al., 2011). In their study, Dehais and colleagues demonstrated that cognitive countermeasures based on temporarily removing fixated information could effectively combat attentional tunneling and help operators disengage from tasks when necessary. This notion of disengagement is precisely consistent with the notion of disengagement in AC (e.g., Shipstead et al., 2016). This technique improved situation awareness by forcing attentional shifts, resulting in better decision making during automation conflicts (Dehais et al., 2011). Such findings suggest that future automation systems could be designed with adaptive interfaces that dynamically modify information presentation based on detected attentional states (or individual differences), thereby supporting the executive control functions of AC while reducing cognitive workload in complex operational environments. This approach (supporting disengagement) would complement existing design principles (e.g. Nielsen, 1994; Norman, 1988; Parasuraman et al., 2000) that focus on supporting relevant information maintenance. Recent work on gaze-contingent displays has addressed attention management from the design side by adapting displays based on operator gaze patterns (e.g., Marois et al., 2020; Taylor et al., 2015). Our findings complement this approach by showing that upstream individual differences in AC, the cognitive ability that produces those gaze patterns, may also warrant consideration. If automation is meant to augment human performance in complex task settings, it must help the user both *maintain* goal-relevant information and *disengage* from irrelevant information.

Limitations and Future Directions

Several limitations warrant consideration. First, out online methodology resulted in data loss due to technical errors and participant dropout. The study required participants to complete six parts over approximately 1.5 h remotely, and our analysis necessitated complete data from each participant. While this methodology has been employed in other large-scale individual difference studies with similar exclusion rates (Burgoyne et al., 2023; Mashburn et al., 2024), and we maintained adequate statistical power despite these exclusions, replication in a more controlled laboratory environment would be beneficial. Second, our between-subjects design with three distinct groups (manual task, information automation, and decision automation) was chosen to reduce participant time commitment. An alternative within-subjects, repeated-measures approach, where participants served as their own controls, would have allowed calculation of within-subject improvement due to automation but would introduce confounds from practice effects and potentially increase dropout rates due to extended study duration. Finally, although the experimental task used has been well-established in previous research, it represents one type of automation (always-on and fixed function allocation) that may not reflect typical real-world automation settings. Future studies should examine the relationship between cognitive abilities and performance with more flexible, adaptable automation (Calhoun, 2022). In such contexts, AC may prove more critical for performance than WM, highlighting the importance of designing automation systems that specifically support AC capabilities.

Key Points

- Working memory requirements were only supported by decision automation, not lower level information automation.
- The study supports the Parasuraman et al. (2000) model, showing that higher levels of automation support working memory demands. However, attention control predicted performance across all conditions, raising questions about whether current automation designs fully accommodate this ability.
- The results of this study indicate a need for automation designs to consider attention demands.

ORCID iDs

Claire Textor https://orcid.org/0000-0002-2831-4621

Richard Pak https://orcid.org/0000-0001-9145-6991

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

References

Bahner, J. E., Hüper, A.-D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, *66*(9), 688–699. https://doi.org/10.1016/j.ijhcs.2008.06.001

Bainbridge, L. (1983). Ironies of automation. *Analysis, Design and Evaluation of Man–Machine Systems*, 129–135.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, *29*(6), 624–630. https://doi.org/10.1177/0963721420969371

Burgoyne, A. P., Mashburn, C. A., Tsukahara, J. S., Pak, R., Coyne, J. T., Foroughi, C., Sibley, C., Drollinger, S. M., & Engle, R. W. (2024). Attention control measures improve the prediction of performance in navy trainees. *International Journal of Selection and Assessment*, *33*(1), e12510. https://doi.org/10.1111/ijsa.12510

Burgoyne, A. P., Pak, R., Draheim, C., Sibley, C., Coyne, J. T., Melick, S. R., & Engle, R. E. (in press). Effects of individual differences in practice on the reliability and validity of the three-minute "squared" cognitive tests. *Journal of Experimental Psychology: General*.

Burgoyne, A. P., Tsukahara, J. S., Mashburn, C. A., Pak, R., & Engle, R. W. (2023). Nature and measurement of attention control. *Journal of Experimental Psychology. General*, *152*(8), 2369–2402. https://doi.org/10.1037/xge0001408

Calhoun, G. (2022). Adaptable (not adaptive) automation: Forefront of human-automation teaming. *Human Factors*, *64*(2), 269–277. https://doi.org/10.1177/00187208211037457

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/10.3758/bf03196772

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain.

Dehais, F., Causse, M., & Tremblay, S. (2011). Mitigation of conflicts with automation: Use of cognitive countermeasures. *Human Factors*, *53*(5), 448–460. https://doi.org/10.1177/0018720811418635

de Visser, E., Shaw, T., Mohamed-Ameen, A., & Parasuraman, R. (2010). Modeling human-automation team performance in networked systems: Individual differences in working

AQ7

memory count. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(14), 1087–1091. https://doi.org/10.1177/154193121005401408

Draheim, C., Pak, R., Draheim, A. A., & Engle, R. W. (2022). The role of attention control in complex real-world tasks. *Psychonomic Bulletin & Review*, *29*(4), 1143–1197. https://doi.org/10.3758/s13423-021-02052-2

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. https://doi.org/10.1037/xge0000783

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492. https://doi.org/10.1080/001401399185595

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43*(2), 226–236. https://doi.org/10.3758/s13421-014-0461-7

Greenwell-Barnden, J. N., Visser, T. A. W., Loft, S., Whitney, S. J., & Bowden, V. K. (2026). How multi-tasking ability impacts performance, workload, situation awareness, stress and trust with simulated imperfect automation. *Human Factors*, *68*(8), 1056–1077. https://doi.org/10.1177/00187208261431968

Huang, F. L. (2022). Reporting results of multilevel models. In *Practical multilevel modeling using R* (pp. 33–48). Sage Publications.

Jamieson, G. A., & Skraaning, G.Jr. (2018). Levels of automation in human factors models for automation design: Why we might consider throwing the baby out with the bathwater. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 42–49. https://doi.org/10.1177/1555343417732856

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*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Koller, M., & Stahel, W. A. (2022). Robust estimation of general linear mixed effects models. In *Robust and multivariate statistical methods: Festschrift in honor of david E. Tyler* (pp. 297–322). Springer International Publishing.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/s0160-2896(05)80012-1

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*(6), 1118–1133. https://doi.org/10.1037/0278-7393.16.6.1118

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Marois, A., Lafond, D., Williot, A., Vachon, F., & Tremblay, S. (2020). Real-time gaze-aware cognitive support system for security surveillance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 1145–1149. https://doi.org/10.1177/1071181320641274

Mashburn, C. A., Burgoyne, A. P., Tsukahara, J. S., Pak, R., Coyne, J. T., Sibley, C., Foroughi, C., & Engle, R. W. (2024). Knowledge, attention, and psychomotor ability: A latent variable approach to understanding individual differences in simulated work performance. *Intelligence*, *104*(■■■), 101835. https://doi.org/10.1016/j.intell.2024.101835

McLaughlin, A. C., & Byrne, V. E. (2020). A fundamental cognitive taxonomy for

cognition aids. *Human Factors*, *62*(6), 865–873. https://doi.org/10.1177/0018720820920099

Miller, C. A., Shaw, T., Emfield, A., Hamell, J., deVisser, E., Parasuraman, R., & Musliner, D. (2011). Delegating to automation: Performance, complacency and bias effects under non-optimal conditions. In *Proceedings of the human factors and ergonomics society annual meeting* (*55*, No. 1, pp. 95–99). Sage Publications.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology. Applied*, *6*(1), 44–58. https://doi.org/10.1037/1076-898x.6.1.44

Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Oberauer, K. (2024). The meaning of attention control. *Psychological Review*, *131*(6), 1509–1526. https://doi.org/10.1037/rev0000514

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

Pak, R., McLaughlin, A. C., & Engle, R. (2024). The relevance of attention control, not working memory, in human factors. *Human Factors*, *66*(5), 1321–1332. https://doi.org/10.1177/00187208231159727

Pak, R., McLaughlin, A. C., Leidheiser, W., & Rovira, E. (2017). The effect of individual differences in working memory in older adults on performance with different degrees of automated technology. *Ergonomics*, *60*(4), 518–532. https://doi.org/10.1080/00140139.2016.1189599

Parasuraman, R., Cosenzo, K. A., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, *21*(2), 270–297. https://doi.org/10.1080/08995600902768800

Parasuraman, R., de Visser, E., Lin, M.-K., & Greenwood, P. M. (2012). Dopamine beta hydroxylase genotype identifies individuals less susceptible to bias in computer-assisted decision making. *PloS One*, *7*(6), e39675. https://doi.org/10.1371/journal.pone.0039675

Parasuraman, R., Kidwell, B., Olmstead, R., Lin, M.-K., Jankord, R., & Greenwood, P. (2014). Interactive effects of the COMT gene and training on individual differences in supervisory control of unmanned vehicles. *Human Factors*, *56*(4), 760–771. https://doi.org/10.1177/0018720813510736

Parasuraman, R., Sheridan, T. B., Wickens, C. D., & New Collective Author (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans: A Publication of the IEEE Systems, Man, and Cybernetics Society*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354. https://www.ncbi.nlm.nih.gov/pubmed/11760769

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL. https://www.R-project.org/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

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

Rovira, E., McGarry, K., & Parasuraman, R. (2007). Effects of imperfect automation on decision making in a simulated command and control task. *Human Factors*, *49*(1), 76–87. https://doi.org/10.1518/001872007779598082

Rovira, E., Pak, R., & McLaughlin, A. (2017). Effects of individual differences in working memory on performance and trust with various degrees of automation. *Theoretical Issues in Ergonomics Science*, *18*(6), 573–591. https://doi.org/10.1080/1463922X.2016.1252806

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110–114.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT press.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*(6), 771–799. https://doi.org/10.1177/1745691616650647

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies*, *51*(5), 991–1006. https://doi.org/10.1006/ijhc.1999.0252

Skraaning, G., & Jamieson, G. A. (2021). Human performance benefits of the automation transparency design principle: Validation and variation. *Human Factors*, *63*(3), 379–401. https://doi.org/10.1177/0018720819887252

Taylor, P., Bilgrien, N., He, Z., & Siegelmann, H. T. (2015). EyeFrame: Real-time memory aid improves human multitasking via domain-general eye tracking procedures. *Frontiers in ICT*, *2*(■■■), Article 17. https://doi.org/10.3389/fict.2015.00017

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

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596x(89)90040-5

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635–654. https://doi.org/10.1080/09658210902998047

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, *8*(3), 201–212. https://doi.org/10.1080/14639220500370105

Wickens, C. D., Helleberg, J., & Xu, X. (2002). Pilot maneuver choice and workload in free flight. *Human Factors*, *44*(2), 171–188. https://doi.org/10.1518/0018720024497943

Wright, M. C., & Kaber, D. B. (2005). Effects of automation of information-processing functions on teamwork. *Human Factors*, *47*(1), 50–66. https://doi.org/10.1518/0018720053653776

Author Biographies

Claire Textor is a human factors researcher at Battelle Memorial Institute. She earned her PhD in human factors psychology from Clemson University in 2023.

Richard Pak is a professor in the Department of Psychology at Clemson University and director of the Clemson Human Factors Institute. He earned his PhD in psychology from the Georgia Institute of Technology in 2005.

Anne Collins McLaughlin is a professor in the Department of Psychology at North Carolina State University. She earned her PhD in psychology from the Georgia Institute of Technology in 2007.

Randall Engle is a professor in the School of Psychology at the Georgia Institute of Technology. He earned his PhD in psychology from the Ohio State University in 1973.