Calibrating Adaptable Automation to Individuals

Jennifer E. Thropp, Tal Oron-Gilad, *Member, IEEE*, James L. Szalma, and Peter A. Hancock, *Fellow, IEEE*

***Abstract*—A detailed understanding of operator individual differences can serve as a foundation for developing a critical window on effective, adaptable, user-centered automation, and even for more autonomous systems. Adaptable automation that functions according to such principles and parameters has many potential benefits in increasing operator trust and acceptance of the automated system. Our current study provides an assessment of the way that individual differences in attentional control (AC) affect the preference for a selection of a desired level of automation (LOA). Participants who scored low or high on AC were either allowed to choose among four possible LOAs or restricted to a predetermined LOA. These manipulations were engaged while the operator was performing visual and auditory target detection tasks. The AC level was found to be inversely proportional to the LOA preference. Operators also performed better when they were preassigned to a fixed LOA rather than given a choice. Individual differences can thus be shown to affect the performance with the automated systems and should be considered in associated design processes. When deciding whether to give the operator control over LOA in a complex system, engineers should consider that the amount of control that operators may want does not necessarily reflect their actual needs.**

***Index Terms*—Human–automation interaction, individual differences.**

I. Introduction

TO CREATE effective, efficient, and a potential for optimal human–machine systems, the affective dimensions of the operator must be assessed. This information then needs to be integrated with the designers' conceptualizations and the engineers' fabrications. Early human–machine systems employed task distribution policies between human and machine that were fixed during design. This static division of function was an offspring of the fractionated time and motion approaches of early manufacturing. However, the shortcomings of such fixed allocation have been exposed across the years, especially as the complexity of systems has increased [1], [2]. In many human–machine systems, there remains a common requirement for humans to exercise supervisory control over even highly automated systems. Here, operators provide correction when necessary and where possible [3]. Even this requirement is now thought to be fading to a greater degree [4].

The evolutionary step that circumvented many of the limitations of static allocation introduced adaptive allocation [2], [5], [6]. Here, the tasking between human and machine could be switched dynamically during operations. This advance was founded on the conception of homeostasis or the dynamic balance that occurs in human bodily functions [2], [7]. Such a strategy looks to alleviate episodes of overload (and underload) by the intelligent redistribution of task demand [8]. This notion has proceeded initially from a purely theoretical proposition (e.g., [9], [10]) to actual implementation [11], [12]. However, adaptive automation is open to further refinement and improvement, especially with respect to augmenting human abilities [13], [14]. The most recent innovations in adaptive allocation have seen the expression of specific operational preferences and the ascendance of "usability" advances. However, a window of opportunity is now opening that can go beyond simple preferences alone to embrace full individual customization [15]. It is to support this latter step in adaptive function allocation that the present investigation is reported. Last, machines will become more automated and autonomous; thus, narrowing the operators' control over them and reducing operators' understanding of how they work or what they do. To approach this transition phase toward more autonomous and semiautonomous systems, there is a need to further study human–automation interaction (e.g., [16]) and human–autonomy interrelations (e.g., [17]).

There is now the feasible practicality of designing human–machine systems that are customized to the unique needs of the individual user. One key construct for trustworthy interaction is control; the degree of perceived control or the need for control may be influenced by personal characteristics. Hence, the purpose of the present paper was to assess one dimension of individual differences, namely attentional control (AC), and to see how variation across this dimension relates to the selection of level of automation through which to achieve a combined set of human–machine tasks. Prior to presenting the experimental aspect of this work, we first consider individual differences in relation to human–machine interaction.

A. Operator Attentional Control in Human–Machine Systems

With user differences potentially accounting for more of the variability in system performance than system design and training variables [18], how individuals respond to different configurations of automation can determine their degree of task success

Manuscript received October 24, 2017; revised April 9, 2018; accepted April 29, 2018. Date of publication June 26, 2018; date of current version November 13, 2018. This work was supported by the Department of Defense Multidisciplinary University Research Initiative (MURI) program, P. A. Hancock, *Principal Investigator*, administered by the Army Research Office under Grant DAAD19-01-1-0621. The views expressed in this work are those of the authors and do not necessarily reflect official Army policy. This paper was recommended by Associate Editor K. M. Feigh. *(Corresponding author: Jennifer E. Thropp.)*

J. E. Thropp is with the School of Graduate Studies, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA (e-mail: throppj@erau.edu).

T. Oron-Gilad is with the Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel (e-mail: orontal@bgu.ac.il).

J. L. Szalma and P. A. Hancock are with the Department of Psychology, University of Central Florida, Orlando, FL 32816, USA (e-mail: james.szalma@ucf.edu; peter.hancock@ucf.edu).

Digital Object Identifier 10.1109/THMS.2018.2844124

or failure. Performance-based studies have revealed that individual response differences among operators as well as less obvious influences such as motivation, emotion, and preferences do influence outcome [19], [20]. An understanding of individual differences can facilitate human–machine interaction [21], [22] possibly through individualized task allocations as driven by operator trait measurements, anticipated error, response behavior, and preferences. The logical inference of this integrated hybrid model is that there are conditions when higher levels of automation (LOAs) do not necessarily reduce workload and improve performance (and see [23]). Different LOAs can instead impose different demands as determined by the unique preferences, behaviors, and abilities of each operator.

In this paper, we address the operator characteristic of AC, which regards the ability to orient one's attention to the appropriate aspects of an environment. It is driven by subfactors for focusing attention and controlling thoughts, which can involve direction of resources to task-relevant stimuli and inhibition of task-irrelevant distractions. Another subfactor of AC is that of shifting attention between tasks, which has implications for multitasking [24], [25]. Attention shifting may affect operator-machine function allocation strategies in automated systems, as attention sharing is a critical component of successful multitask performance [26] and it is likely that individuals vary in the resources they can allocate and the efficiency of such allocation. Specifically, for the current study, we chose to focus on differences in operators' AC within the contexts of various LOAs and task demands.

As attention is a limited resource [27], performance declines as these resources are allocated to multiple tasks at the same time. Should one of the concurrent tasks become more automated, however, performance on the manually controlled task may rebound as resources have been freed to serve it [28]. Various LOAs can therefore affect task difficulty, allowing for improved dual-task performance as automation assumes more control. An operator with good attentional allocation may thrive at the helm of simultaneous inputs and duties, thereby manually controlling multiple system functions effectively. Conversely, an operator with poor allocation skills may have difficulty managing such parallel demands and prefer data to be presented in a singular mode or only on specific task components. With the assistance of higher LOAs, however, low AC operators may succeed as task demands increase.

In addition, there may be interactive AC and LOA implications based on modality. Auditory signal detection has shown superiority over visual target detection in terms of speed, accuracy, and resistance to a vigilance decrement [29], [30], likely in part because auditory stimuli and auditory perception are closely coupled [31]. Auditory processing is omnidirectional by nature in that sound can be received from any direction, whereas visual processing involves selective attention to scan for information [32]. There is thus looser coupling between visual stimuli and visual perception, as the observer can avert his or her eyes and head away from the stimulus, potentially missing targeted information. Poor attentional allocation thus may be especially adverse for extensive visual task demands. Here, adaptable systems could permit parsing information and tasks into separate displays or different modalities within the interface and the operator could delegate more control to the system in visually oriented tasks.

B. Experimental Hypotheses

When given the choice among a range of LOAs, it was hypothesized that participants with low AC would exhibit a general preference for higher LOAs. Relative to when in lower LOAs, their behavior in higher LOAs would reflect the following:

1) reduced subjective workload;
2) improved dual-task performance;
3) improved perceptual sensitivity across all tasks;
4) a response bias more like that of the automation, which in the current study, was programmed to have a neutral policy.

Participants with high AC, on the other hand, were hypothesized to prefer lower LOAs. This adoption of lower LOAs would attenuate the influence of the automation's neutral response policy, and thus their own innate biases would prevail. Regardless of LOA, it was anticipated that they would generally exhibit better task performance than those with low AC, with higher sensitivity and generally lower workload across tasks.

In the participants' initial screening phase, we assessed the distribution of AC within the elicited sample and distinguished subsequent participants as either low- or high-scoring on the AC scale (see [24]). We then assessed if there was a relationship between AC scores and self-reported preference for automation level in a hypothetical scenario. The first experiment sampled individuals from the screening phase who showed either low or high AC scores and examined their preferences for various LOAs, performance on target detection tasks, and subjective workload during these tasks. The second experiment also tested performance and subjective workload in low and high AC groups; however, participants were assigned to low or high LOAs rather than being allowed a choice.

II. Experiment 1—Low and High Levels of AC in Adaptable Levels of Automation

A. Experimental Participants and Tasks

Three hundred and three undergraduates (227 females and 76 males) were recruited from undergraduate psychology courses at a U.S. university to create an initial pool of participants. Their ages ranged from 18 to 39 years (M = 22.3, SD = 3.5). AC was assessed using the AC scale [24], in which scores could range from 20 (low) to 80 (high). Across all sampled participants, AC scores ranged from 24 to 74 (M = 49.6, SD = 9.2). Based upon visual inspection, the AC scores were normally distributed with neither obvious skewness nor kurtosis (see [33]). To separate low and high AC scores and maximize variance between the groups, a quartile split was performed; the high AC category was determined by scores ranging from 56 to 74, whereas the low AC scores ranged from 24 to 44.

1) Automation Level Preference Index: To achieve a preliminary general assessment of preference for LOA, the initial sample (n = 303) completed the automation level preference index. After reading a hypothetical scenario, they indicated the percentage of computerized assistance they preferred in a novel target detection-type task, in which the computer made correct detections 90% of the time (i.e., 90% reliability). They responded using a scale of 0% (preference for no computer assistance whatsoever; i.e., a fully manual task) to 100% (preference for full computer control; i.e., a fully automated task).

Scores ranged from 11% to 90% ($M = 67.5$, SD $= 16.4$) among those with low AC, and from 0% to 88% ($M = 61.5$, SD $= 20.6$) among those with high AC. An independent samples t-test indicated that participants with low AC scores preferred higher levels of intervention compared to those with high AC scores, $t(146.974) = 2.03$, $p < 0.05$, Cohen's $d = 0.32$. While both groups generally preferred greater intervention, it was anticipated that AC may have implications for LOA preferences in the subsequent target detection task experiment.

2) Target Detection Tasks: From the initial pool of 303 participants, 18 participants with scores in the low AC range and 18 participants with scores in the high AC range were selected. Among those selected, scores in the low AC group ranged from 36 to 44 ($M = 40$, SD $= 2.5$) and scores in the high AC group ranged from 56 to 73 ($M = 60.25$, SD $= 4.2$). Participants (28 females, 8 males) ranged in age from 18 to 34 years ($M = 22.2$, SD $= 3.5$). The ages and ratio of female to male participants were representative of the study-eligible population of undergraduate psychology students. Previous work has indicated little effect in terms of gender differences in temporally oriented signal detection tasks such as the one employed herein (e.g., [34]). Furthermore, automated systems are not at present designed differently for male and female operators.

The experimental target detection tasks consisted of visual and auditory stimuli, with five different LOAs configured to assist the participant. The visual stimuli were 32 mm black vertical lines appearing on the left and right sides of a white 15.5-inch computer display. Target lines appeared for a short (125 ms) duration; nontarget lines appeared for a long (250 ms) duration. The auditory stimuli were tones (400 Hz C4 triangle waveform; see [35]) sounding in the right and left earphones separately. Target tones sounded for a short duration (200 ms); nontarget tones sounded for a long duration (250 ms). There were thus four stimulus types that were presented singularly in a random sequence during each task, with the restriction that no two stimuli were presented simultaneously.

A temporally oriented discrimination was chosen because stimulus presentation time is a characteristic that can be manipulated in both auditory and visual displays. The temporal values used to distinguish target from nontarget stimuli in both modalities were psychophysically equated for discriminability (see [29]). The difference between the target and nontarget stimuli was greater in the visual task than in the auditory task (i.e., 125 ms versus 50 ms delta, respectively) to compensate for the finding that temporal perception in the auditory modality is more sensitive than in the visual modality (see [36]). Additionally, before undergoing a task training session, each participant matched the apparent loudness of the tones to the apparent brightness of the display using a cross-modality matching technique (see [37]). This matching was intended to psychophysically equate the perceptual intensity of both visual and auditory stimuli, thereby precluding any bias resulting from greater salience of stimuli of either modality.

TABLE I
NUMBER OF EVENTS PER TASK AT 90% MACHINE RELIABILITY

Task	Hits	Correct Rejections	Misses	False Alarms	Total
*Single-Task	18	18	2	2	40
**Dual-Task	36	36	4	4	80

*Auditory-Only or Visual-Only Tasks **Audiovisual Combined Task

The automation (i.e., "machine") identified stimuli and the participant could observe its decisions. Its decision was presented to the participant using a dialog box with either the word "Target" or "Non" appearing for 3 s after each visual stimulus, or the word "Target" or "Non" sounding in the earphones after each auditory stimulus. The term "non" was used instead of "nontarget" to increase the verbal distinction from the term "target."

Each task consisted of 80 events (40 visual and 40 auditory) and lasted approximately 80 min. The automation was programmed to operate at a reliability of 90% in terms of both target and nontarget detection assistance, and thus execute a specified number of hits, correct rejections, misses, and false alarms (see Table I). A high reliability was used here as encouragement for participants to trust and use the automation as they deemed necessary, as reliability affects the extent to which operators trust it [38], [39]. A 90% reliability has been used as a high reliability condition in previous studies [39], [40].

The three tasks were as follows:

1) *Visual-Only Task (Single-Task):* Participants differentiated between target and nontarget lines using their choice of LOA. Meanwhile, the concurrent auditory detection task was fully automated. The participant could monitor the automation and click on the "Reset" button if she/he noticed it making incorrect decisions but could not change its decisions.
2) *Auditory-Only Task (Single-Task)*: Participants distinguished between target and nontarget sounds using their choice of LOA. Meanwhile, the concurrent visual detection task was fully automated. As in the visual-only task, the participant could monitor the automation's decisions and click "Reset" if thought necessary but could not change its decisions.
3) *Audiovisual Combined Task (Dual-Task):* Participants were responsible for executing both the auditory and visual target detection tasks simultaneously and could choose the LOA for the auditory and visual tasks separately. They could not choose the fully automated level for either task.

B. LOA Algorithms

The automation application offered five different LOAs following specific operator-machine response algorithms:

LOA 5. Fully Machine-Controlled: The machine identified all lines while the participant performed the auditory-only task, or all tones while the participant performed the visual-only task. The participant could monitor the machine and click "Reset," but not change its decisions.

LOA 4. Mostly Machine-Controlled: The machine was responsible for identifying all lines in the visual tasks (i.e., visual-only and visual portion of the audiovisual combined tasks) or all tones in the auditory tasks (i.e., auditory-only and auditory portion of audiovisual combined tasks). The participant could monitor the machine's decisions and had a 3-s window in which to change each machine response by clicking on "Veto," with the goal of exceeding its 90% correct detection rate. The participant could also opt to indicate agreement with the machine responses by clicking on "Agree," although doing so did not change the machine's decisions. If the participant did not veto any of the machine responses, his/her accuracy for that task would be 90%.

LOA 3. Equal Sharing (Half Operator Control, Half Machine Control): The system used LOA 2 (Mostly Operator-Controlled as described below) for lines appearing on the left side of screen and tones sounding in the left earphone, and LOA 4 (Mostly Machine-Controlled) for lines appearing on the right side of screen and tones sounding in the right earphone. This arrangement is a common work distribution in manned–unmanned teams (MUM-T).

LOA 2. Mostly Operator-Controlled: The participant was responsible for identifying all lines in the visual tasks, or all tones in the auditory tasks, and did so by clicking on "Target" on "Non" after each stimulus was presented. Then, the machine made an independent decision with 90% reliability. The participant could then change his/her decision by clicking on "Target" on "Non" as appropriate.

LOA 1. Fully Operator-Controlled. The participant identified all lines in the visual task, or all tones in the auditory task. The machine did not identify any stimuli.

C. Choosing LOAs

Participants could select only among LOAs 1 through 4 by clicking on the options on the control mode menu on the screen (see Fig. 1) and could change LOAs as often as desired throughout the tasks. In the visual-only task, participants selected their preferred LOA only for visual stimulus detection, while the concurrent auditory task was fixed at LOA 5 (i.e., fully automated). Conversely, in the auditory-only task, participants selected their preferred LOA only for auditory stimulus detection, while the concurrent visual task was fixed at LOA 5. In the dual-task audiovisual combined task, participants selected their preferred LOAs for both the visual and auditory tasks, but could not select LOA 5 for either one, as it was only used in the task that was fully automated during the two single-task conditions.

D. Experimental Design

A 4 (selectable LOA: Fully Manual/Mostly Manual/Equal Manual-Machine/Mostly Machine) × 2 (Modality: Auditory/ Visual) × 2 (Load: Single-Task/Dual-Task) within-participants experimental design was employed. AC group (low and high) was the between-participants variable. Participants performed the three target detection tasks and were instructed to select the LOAs that they felt would allow them to best complete each task. Fig. 1 shows the automation interface.

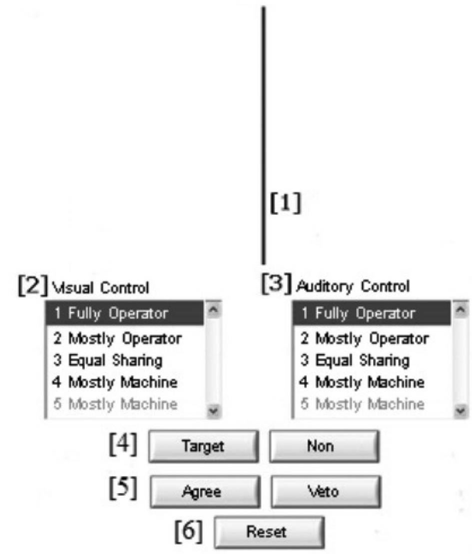


Fig. 1. Automation interface. The vertical division line [1] determined left and right sides of the screen. Participants could choose the LOA for the visual [2] and auditory [3] stimuli; the current LOAs are highlighted. They clicked "Target" or "Non" [4] based on their decision for each stimulus, "Agree" or "Veto" [5] after each machine decision (optional), and "Reset" [6] if they suspected an automation failure (also optional).

E. Dependent Measures

Participants' preferences for the different LOAs were assessed by the proportion of time that they chose to be in each of the LOAs relative to the entire duration of the task. The number of times participants opted to click "Agree" or "Veto" was recorded, as doing so may indicate their system interaction behavior and attentiveness to the machine's input.

Responses were analyzed using the principal elements of signal detection theory (see [41]), particularly the derived measure d' which is an assessment of the participants' sensitivity to the distinction between targets and nontargets, and β which is a measure of the response bias toward conservatism or leniency in reporting stimuli as targets [42]. A high d' indicates that the participant was more likely to correctly discriminate between targets and nontargets. A high β indicates that the participant was more likely to categorize stimuli as nontargets and commit more misses, suggesting a more cautious response criterion. Conversely, a low β indicates a greater likelihood to classify stimuli as targets and commit more false alarms, suggesting a more lenient response criterion.

1) Load Effects: Four separate task components were analyzed. The visual-only and auditory-only tasks were single-task conditions. The visual portion (i.e., "visual-combined") and the auditory portion (i.e., "auditory-combined") of the audiovisual combined task were both part of the dual-task condition. Thus, two task load conditions were compared: single and dual.

2) Modality Effects: Modality effects were assessed across their respective tasks, such that the visual modality effects encompass both the visual-only and visual-combined tasks, whereas the auditory modality effects encompass both the auditory-only and auditory-combined tasks.

3) Machine Effects: The responses of the operator were analyzed in relation to the machine's decisions (i.e., without and with machine aid). Operator-alone responses were the participants' responses independent of the machine aid, whereas machine-aided responses were the participants' final responses after incorporating the machine aid. Thus, there were two levels of response: operator-alone and machine-aided.

4) Subjective Workload Assessment: Perceived workload was assessed using the abbreviated version of the Raw NASA Task Load Index (RTLX), which uses an unweighted average of six subscale values. In the original version of the TLX [43], paired comparisons were used to derive weights for the six subscales. However, as Byers, Bittner, and Hill in [44] have shown and as discussed by Nygren [45], RTLX scores can provide a better account of the workload experienced. Participants rated their perceived workload on the individual subscales on a scale from 0 to 100, with higher values indicating higher workload. Global RTLX estimates were derived by calculating the average of the subscale values.

F. Experimental Procedure

Participants first completed a task training session in which they were presented with task instructions and demonstrations of each LOA in use, and then practiced each of the three tasks for two minutes at each LOA using the automation application.

In the experimental sessions, the order of the two single-task conditions was counterbalanced across participants and the audiovisual-combined task was completed last. Participants were instructed to choose from among LOAs 1 through 4 and change LOAs throughout the tasks as necessary to maximize their task performance. After each task, they completed the abbreviated version of the RTLX.

III. Experiment 1 Results

A. Preferences for LOAs

A 2 (AC) $\times$ 2 (Modality) $\times$ 2 (Load) $\times$ 4 (LOA) mixed design ANOVA was conducted with repeated measures on the last three factors. There was a significant main effect for LOA, $F(3, 105) = 6.84, p < 0.0001$. A post hoc test conducted using Tukey's HSD ($q = 3.83$, $t = 2.70$) revealed that compared to LOA 3, significantly more time was spent at LOA 1, $t(35) = 3.26, p < 0.005$, LOA 2, $t(35) = 5.26, p < 0.0001$, and LOA 4, $t(35) = 4.86$, $p < 0.0001$. The proportion of time spent at each LOA is shown in Table II.

The main effect for LOA was predominantly driven by the result that very little time was spent at LOA 3, leading to a significant Mauchly's test of sphericity for LOA and all interactions among it and other variables (i.e., modality by LOA, load by LOA, modality by load by LOA; all $p < 0.005$). This effect was thought to mask effects seen with the other three LOAs; therefore, LOA 3 was removed from the ANOVA.

All further analyses were conducted considering only LOAs 1, 2, and 4. The revised experimental design was thus a 2 (Modality) $\times$ 2 (Load) $\times$ 3 (LOA) repeated measures ANOVA with AC as a between-participants variable. There was no significant main effect for LOA, $F(2, 68) = 1.43, p > 0.05$.

TABLE II
Mean Proportion of Time Spent at Each LOA

LOA	Overall Mean (SD)	Auditory Tasks Mean (SD)	Visual Tasks Mean (SD)
1: Full Operator	.218 (.35)	.244 (.37)	.191 (.35)
2: Mostly Operator	.397 (.42)	.441 (.44)	.351 (.44)
3: Equal Sharing	.024 (.07)	.030 (.11)	.019 (.08)
4: Mostly Machine	.361 (.41)	.285 (.46)	.439 (.43)

There was a significant interaction between modality and LOA, $F(2, 68) = 10.34, p < 0.0001$. A post hoc paired samples t-test showed that in the auditory tasks, operators spent more time at LOA 2 than they did in the visual tasks, $t(35) = 2.39$, $p < 0.05$, and vice versa. In the visual tasks, LOA 4 was used more often than it was in the auditory tasks, $t(35) = 2.74$, $p = 0.01$. Only in the visual modality was the total time at LOA 4 significantly higher than the time spent at LOA 1, $t(35) = 2.31$, $p < 0.05$. Furthermore, post hoc paired-sampled t-tests showed that use of LOA 4 in the dual-task conditions (across modalities) was significantly higher than in the single task conditions ($M = 0.41$, SD $= 0.43$ versus $M = 0.30$, SD $= 0.39$, respectively), $t(35) = 2.25, p < 0.05$, Cohen's $d = 0.26$.

B. AC Levels and LOA Preferences

Although there was no significant interaction between LOA and AC, $F(2, 68) = 0.59$, $p > 0.05$, there was a trend for those with low AC to use higher LOAs than those with high AC across all four tasks [see Fig. 2(a)]. A preplanned test of simple effects indicated that those with low AC spent significantly more time at LOA 4 ($M = 0.48$, SD $= 0.47$) than LOA 1 ($M = 0.10$, SD $= 0.23$), $t(17) = 2.85, p < 0.05$, Cohen's $d = 1.17$. There was a trend for those with low AC ($M = 0.44$, SD $= 0.38$) to spend more time at LOA 4 than those with high AC ($M = 0.29$, SD $= 0.45$), Cohen's $d = 0.56$. Conversely, those with high AC ($M = 0.30$, SD $= 0.42$) spent more time at LOA 1 than those with low AC ($M = 0.14$, SD $= 0.24$), Cohen's $d = 0.56$. There was a significant interaction between LOA and modality, $F(2, 68) = 3.31, p < 0.05$; in visual tasks, low AC participants spent significantly more time at LOA 4 ($M = 0.57$, SD $= 0.43$) than LOA 1 ($M = 0.08$, SD $= 0.23$). Fig. 2(b) illustrates the proportion of time spent at each LOA per task for the low and high AC groups.

C. Signal Detection Performance – d' (Sensitivity)

A 2 (AC) $\times$ 2 (Modality) $\times$ 2 (Load) $\times$ 2 (without-with Machine Aid) mixed design ANOVA was computed for perceptual sensitivity (d') with repeated measures on the last three factors. Two main effects, modality, $F(1, 34) = 52.34, p < 0.0001$, and machine aid, $F(1, 34) = 4.65, p < 0.05$, were statistically significant. The three-way interaction of AC by modality by load was also significant, $F(1, 34) = 5.62, p < 0.05$ [see Fig. 3(a)].

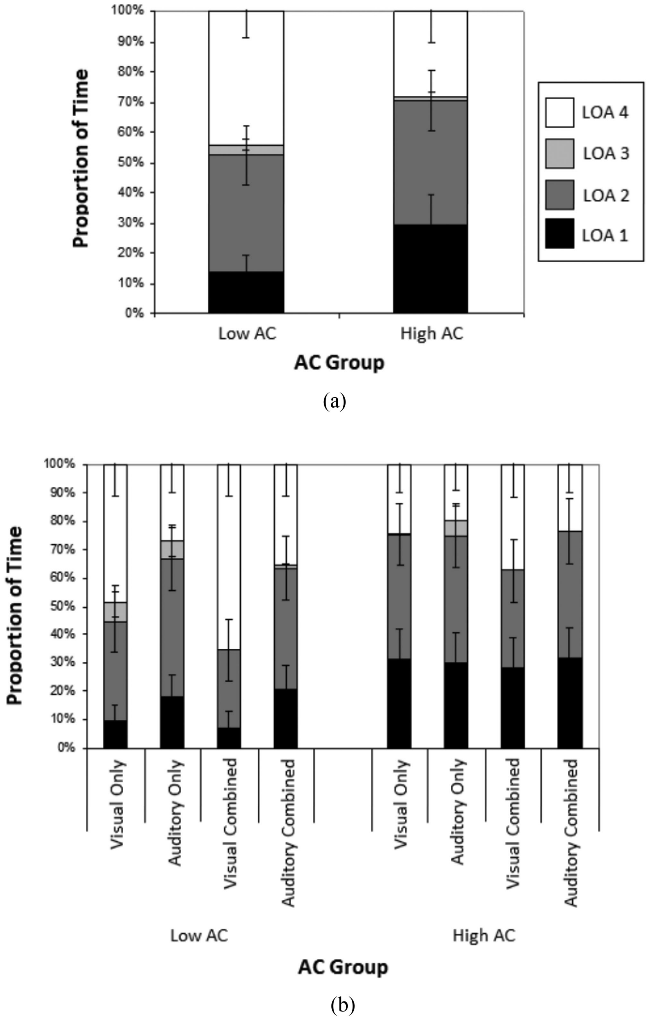


Fig. 2. (a) Overall proportion of time spent at each LOA across tasks for low and high AC groups in Experiment 1. (b) Proportion of time spent at each LOA for each task for low and high AC groups in Experiment 1.

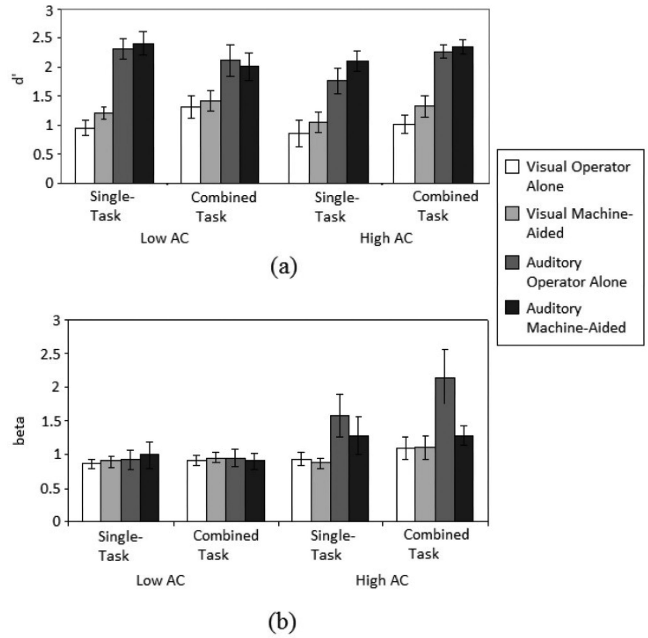


Fig. 3. (a) Operator-alone and machine-aided d' in each task grouped by AC in Experiment 1. (b) Operator-alone and machine-aided β in each task grouped by AC in Experiment 1.

The d' in the auditory task $(M = 2.16, \mathrm{SD} = 0.57)$ was significantly higher than in the visual task $(M = 1.14, \mathrm{SD} = 0.55)$, Cohen's $d = 1.82$. The use of the machine aid improved operators' sensitivity across tasks and modalities. Specifically, in the audiovisual-combined task, there was a significant effect for use of the machine aid, $t(35) = 2.14$, $p < 0.05$, such that in the visual-combined task, the operators' mean d' was significantly higher when the machine aid was used $(M = 1.25, \mathrm{SD} = 0.59$ versus $M = 1.04, \mathrm{SD} = 0.66)$, Cohen's $d = 0.34$. There were no significant improvements in d' through using the machine aid in either of the single-task conditions.

For the high AC group, the overall d' in the auditory-combined task $(M = 2.31, \mathrm{SD} = 0.76)$ was significantly higher than that of visual-combined task $(M = 1.17, \mathrm{SD} = 0.62)$, $t(17) = 5.90$, $p < 0.0001$, Cohen's $d = 0.36$, and the auditory-only task $(M = 1.93, \mathrm{SD} = 0.45)$, $t(17) = 2.18$, $p < 0.05$, Cohen's $d = 0.71$. Also, their overall d' in the auditory-only task was significantly higher than in the visual-only task $(M = 0.95, \mathrm{SD} = 0.61)$, $t(17) = 4.12$, $p = 0.001$, Cohen's $d = 1.81$.

For participants with low AC, the overall d' in the auditory-only task $(M = 2.36, \mathrm{SD} = 0.77)$ was significantly higher than in the visual-only task $(M = 1.08, \mathrm{SD} = 0.45)$, $t(17) = 6.96$, $p < 0.0001$, Cohen's $d = 2.36$. Their overall d' in the auditory-combined task $(M = 2.06, \mathrm{SD} = 1.06)$ was also significantly higher than in the visual-combined task $(M = 1.37, \mathrm{SD} = 0.76)$, $t(17) = 2.41, p < 0.05$, Cohen's $d = 0.92$.

D. Signal Detection Performance—β (Response Bias)

A 2 (AC) × 2 (Modality) × 2 (Load) × 2 (without–with Machine Aid) mixed design ANOVA was computed for response bias (β), with repeated measures on the last three factors. Main effects were significant for AC, $F(1, 34) = 8.19$, $p < 0.01$, modality, $F(1, 34) = 8.448$, $p < 0.01$, and machine aid use, $F(1, 34) = 4.15$, $p < 0.05$. Two-way interactions were significant for modality by AC, $F(1, 34) = 6.41$, $p < 0.02$, and machine aid use by AC, $F(1, 34) = 5.76, p < 0.05$.

In the auditory modality, β was significantly higher (i.e., more conservative response bias) than in the visual modality $(M = 1.26, \mathrm{SD} = 0.67$ versus $M = 0.95, \mathrm{SD} = 0.37)$. Altogether, participants with high AC $(M = 1.29, \mathrm{SD} = 0.49)$ had a significantly higher mean β than those with low AC $(M = 0.92, \mathrm{SD} = 0.23$, Cohen's $d = 0.49)$. More specifically, those with high AC $(M = 1.44, \mathrm{SD} = 0.62)$ had a significantly higher mean operator-alone β than those with low AC $(M = 0.91, \mathrm{SD} = 0.24)$, $t(34) = 3.36$, $p < 0.005$, Cohen's $d = 1.20$. There was no significant difference in machine-aided β between those with low and high AC. Fig. 3(b) shows the mean β values for each task and condition of machine use for the low and high AC groups.

E. Subjective Workload Ratings

A two-way (Task) repeated measures ANOVA was conducted. AC was the between-participants variable. There was a significant main effect for task, $F(2, 68) = 9.91, p < 0.005$. The combined task ($M = 51.1$, SD $= 13.6$) had significantly higher overall workload ratings than the auditory-only task ($M = 45.5$, SD $= 9.8$), $t(35) = 3.34$, $p < 0.005$, Cohen's $d = 0.47$, and the visual-only task ($M = 46.7$, SD $= 13.2$), $t(35) = 3.19, p < 0.005$, Cohen's $d = 0.32$.

A priori tests of simple effects found that the high AC group ($M = 51.4$, SD $= 12.0$) reported significantly higher overall workload in the visual-only task than the low AC group ($M = 42.0$, SD $= 12.9$), $t(34) = 2.27, p < 0.05$, Cohen's $d = 1.11$.

F. Human–Machine Interaction

Table III specifies the mean number of LOA changes for each type of task condition. There was no significant effect for AC group, although there was a trend for low AC participants to change LOAs more often.

Participants were generally more interactive with the automation in auditory tasks than in visual tasks; the mean number of times each participant opted to click on "Agree" or "Veto" in LOA 4 was significantly higher in auditory tasks ($M = 12.00$, SD $= 14.74$) than visual tasks ($M = 7.97$, SD $= 12.02$), $t(35) = 2.62$, $p < 0.05$. Although there was no significant main effect for AC, $F(1, 34) = 1.18, p = 0.192$, there was a trend for high AC participants ($M = 12.81$) to respond more often than those with low AC ($M = 7.26$), Cohen's $d = 0.53$.

IV. Experiment 1 Discussion

Statistically, preferences for LOAs 1, 2, and 4 did not differ across tasks or levels of AC. LOA 2 was used extensively in both AC groups, perhaps since machine suggestions provided a type of reliable performance feedback. Comments from participants of both groups indicated that, first, the tasks were challenging and, second, whereas they generally trusted their own decisions, receiving suggestions from a largely accurate decision aid was reassuring. Both groups spent minimal time at LOA 3, which used a combination of LOAs 2 and 4. This arrangement was not favorable, possibly due to increased effort to develop response strategies. In previous studies, Tulga and Sheridan in [46] and Olson and Sarter in [10] also observed that the effort in developing strategies for task management serves to increase workload.

Low AC participants spent significantly more time in LOA 4 than LOA 1, but only in visual tasks. This result may be due to the weaker coupling between visual perception and visual stimuli compared to that of the auditory modality [31], which could consequently further disadvantage low AC operators, especially without automated assistance.

Incorporation of the machine aid was more effective in regulating β for those with high AC. These individuals may have been superior in their ability to generate their own responses as well as consider those of the machine, as doing so would demand increased attentional manipulation.

TABLE III
MEAN NUMBER OF LOA CHANGES DURING EACH TASK

Condition	All Mean (SD)	Low AC Mean (SD)	High AC Mean (SD)
Visual-Only	1.14 (1.13)	1.28 (1.07)	1.00 (1.19)
Auditory-Only	.94 (0.63)	1.06 (0.64)	.83 (0.62)
Visual-Combined	.86 (0.64)	.94 (0.73)	.78 (0.65)
Auditory-Combined	1.14 (1.33)	1.39 (1.65)	.89 (0.90)

Participants of both AC groups clicked on "Agree" or "Veto" several times per task, a trend that is beneficial for systems that learn user preferences through interaction. The high AC group, however, was generally more interactive. Further research could examine why low AC participants responded less often (e.g., if less interested or less confident) and how engagement could be encouraged.

The high AC group reported significantly higher overall workload in the visual-only task than the low AC group. Though this result may contradict the hypothesis that those with high AC would report lower workload values, it may reflect the fact that they spent more time in LOA 1 than the low AC group. Further, higher workload may not necessarily imply overload, but may indicate a satisfactory level of task engagement.

V. Experiment 2—Low and High AC in Fixed Levels of Automation

Given the discrepancy between automation need and preference [47], it cannot be determined if the LOAs chosen in Experiment 1 truly optimized task performance and workload, especially with respect to individual differences in attentional behavior. With LOA flexibility versus rigidity a consideration in operator-machine system design, a second experiment was conducted in which low and high AC participants did not have a choice in LOA but were instead preassigned to either a low or high LOA that could not be changed.

From the same initial pool, 32 new participants (21 females, 11 males) with AC scores in either the low or high ranges were selected. The low AC group ($n = 16$) had AC scores ranging from 34 to 43 ($M = 39.7$, SD $= 2.7$) and the high AC group ($n = 16$) had AC scores ranging from 56 to 67 ($M = 61.3$, SD $= 2.6$). They ranged in age from 18 to 34 years ($M = 21.2$, SD $= 2.7$) and were representative of the sampled population of undergraduate psychology students. Participants from the low and high AC groups were equally and randomly assigned to either of two fixed LOAs: 2 (Mostly Operator-Controlled) or 4 (Mostly Machine-Controlled). LOAs 2 and 4 were utilized because they both provided the opportunity to observe a machine-aided component of operator performance. The automation algorithms and task requirements were identical to those in Experiment 1; the only exception was that participants were not able to change their LOAs.

VI. Experiment 2 Results

A. Signal Detection Performance—d' (Sensitivity)

A 2 (AC) $\times$ 2 (LOA) $\times$ 2 (Modality) $\times$ 2 (Load) mixed design ANOVA was conducted with repeated measures on the last two factors. The machine answers were the default responses of participants in LOA 4 unless they opted to veto the machine responses; as such, LOA 4 only had a machine-aided component and did not have an operator-alone component. LOA 2 had both operator-alone and machine-aided components since the operator answer was the default but could be influenced by machine suggestions. Thus, to produce a fair comparison between LOAs 2 and 4, only the machine-aided d' from LOA 2 was compared to the machine-aided d' from LOA 4. Significant main effects were found for LOA, $F(1, 28) = 4.58$, $p < 0.05$, and modality, $F(1, 28) = 144.12$, $p < 0.0001$. No interactions were significant and there was no main effect for AC.

LOA 4 had higher d' values ($M = 2.27$, SD $= 0.38$) than LOA 2 ($M = 1.77$, SD $= 0.69$), Cohen's $d = 0.90$. Further, the auditory machine-aided d' ($M = 2.84$, SD $= 0.59$) was significantly higher than the visual machine-aided d' ($M = 1.37$, SD $= 0.53$), Cohen's $d = 2.62$. A separate set of *a priori* tests of simple effects found that in the visual-only task in LOA 2, the machine-aided d' ($M = 1.20$, SD $= 0.65$) was significantly higher than the operator-alone d' ($M = 0.50$, SD $= 0.60$), $t(15) = 2.98$, $p < 0.01$, Cohen's $d = 1.15$. Table IV provides descriptive statistics for d'.

TABLE IV
SENSITIVITY (d') IN THE FIXED LOA CONDITIONS OF EXPERIMENT 2

	LOA 2 Mean (SD)		LOA 4 Mean (SD)	
Condition	Low AC	High AC	Low AC	High AC
Visual-Only	1.45 (.57)	0.95 (.66)	1.56 (.48)	1.49 (.62)
Auditory-Only	2.68 (.65)	2.65 (.95)	2.75 (.51)	2.87 (.52)
Visual-Combined	1.30 (.46)	1.12 (.78)	1.53 (.56)	1.57 (.38)
Auditory-Combined	2.67 (.81)	2.69 (.60)	3.31 (.41)	3.06 (.76)

B. Signal Detection Performance—β (Response Bias)

A 2 (AC) × 2 (LOA) × 2 (Modality) × 2 (Load) mixed design ANOVA was conducted with repeated measures on the last two factors. As with the *d*' analyses, the ANOVA compared only the final response biases for each LOA group: the machine-aided β for LOA 2 and the machine-aided β for LOA 4. Significant effects were found for AC, $F(1, 28) = 6.82$, $p < 0.05$, and the interaction of modality by AC, $F(1, 28) = 10.35$, $p < 0.005$; this latter interaction is illustrated in Fig. 4.

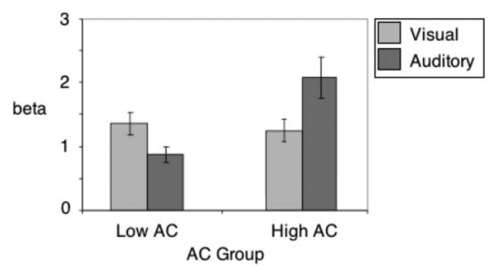


Fig. 4. Mean machine-aided β by AC group and modality in Experiment 2.

Participants with high AC ($M = 1.66$, SD $= 0.73$) had a higher overall β than those with low AC ($M = 1.11$, SD $= 0.45$), Cohen's $d = 1.07$, and a higher β for the auditory modality ($M = 2.07$, SD $= 1.28$) than for the visual modality ($M = 1.25$, SD $= 0.70$), $t(15) = 2.25$, $p < 0.05$, Cohen's $d = 0.91$. Individuals with low AC, however, had a higher β for the visual modality ($M = 1.36$, SD $= 0.71$) than for the auditory modality ($M = 0.87$, SD $= 0.51$), $t(15) = 2.35$, $p < 0.05$, Cohen's $d = 0.98$.

In LOA 2, tests of simple effects showed that in the auditory-only task, the high AC group's machine-aided β ($M = 1.62$, SD $= 1.27$) was significantly higher than that of the low AC group ($M = 0.54$, SD $= 0.39$), $t(14) = 2.29$, $p < 0.05$, Cohen's $d = 1.20$.

C. Subjective Workload Ratings

There was a significant main effect for task, $F(2, 28) = 4.56$, $p < 0.05$. The visual-only task had significantly higher workload ratings than the auditory-only task, $t(31) = 2.56$, $p < 0.05$, Cohen's $d = 0.28$. The combined task had significantly higher workload ratings than the visual-only task, $t(31) = 2.18$, $p < 0.05$, Cohen's $d = 0.18$, as well as the auditory-only task, $t(31) = 3.32$, $p < 0.005$, Cohen's $d = 0.46$. There was no significant main effect for LOA or AC, yet there was a trend for high AC participants to report higher workload ratings. See Table V for the subjective workload ratings. There was also a trend for higher workload in LOA 2 than LOA 4 for the low AC group in the visual-only ($M = 45.4$, SD $= 24.3$ versus $M = 39.6$, SD $= 16.3$) and auditory-only tasks ($M = 39.7$, SD $= 19.2$ versus $M = 33.9$, SD $= 9.6$). Conversely, there was a trend for higher workload in LOA 4 than LOA 2 for the high AC group in the visual-only ($M = 49.1$, SD $= 10.7$ versus $M = 46.6$, SD $= 15.6$) and auditory-only tasks ($M = 45.3$, SD $= 23.2$ versus $M = 42.8$, SD $= 16$).

TABLE V
MEAN WORKLOAD RATINGS IN FIXED LOAS OF EXPERIMENT 2

Condition	Overall (SD)	LOA 2 (SD)	LOA 4 (SD)	Low AC	High AC
Visual-Only	45.2 (16.9)	46.0 (19.7)	44.4 (14.2)	42.5 (20.2)	47.8 (13.0)
Auditory-Only	40.4 (17.4)	41.2 (17.2)	39.6 (18.2)	36.8 (15.0)	44.1 (19.3)
Audiovisual Combined	48.5 (19.3)	50.3 (23.3)	46.7 (14.7)	45.5 (21.4)	51.6 (17.0)

D. Human–Machine Interaction

There was a significant interaction between modality and LOA group in the mean number of times participants clicked on "Agree' or "Veto," $F(1, 28) = 27.25$, $p < 0.0001$. In LOA 2, participants did so in the auditory modality ($M = 23.5$, SD $= 15.2$) more than in the visual modality ($M = 17.2$, SD $= 11.5$), $t(15) = 2.95$, $p = 0.01$, Cohen's $d = 0.47$. The pattern was reversed in LOA 4, however, with participants responding to the machine more often in the visual modality ($M = 12.3$, SD $= 4.8$) than the auditory modality ($M = 5.3$, SD $= 2.8$), $t(15) = 5.673$, $p < 0.0001$, Cohen's $d = 1.79$.

VII. EXPERIMENT 2 DISCUSSION

As hypothesized, the use of the higher LOA improved performance for both low and high AC participants; the machine-aided

d' was higher in LOA 4 than in LOA 2. An interesting interaction between AC and modality was observed; the mean β for low AC participants was higher in the visual modality, but the mean β for the high AC participants was higher in the auditory modality.

The workload patterns for each task also closely replicated the results from the first experiment, such that the combined task typically had the highest workload ratings, and the visual-only task workload exceeded that of the auditory-only task. The demand of each task was, therefore, subjectively similar across both adaptable and fixed LOAs. As hypothesized, using LOA 2 yielded higher workload ratings than LOA 4 across AC groups. There was also evidence for reduced workload at LOAs that were more complementary to AC level, notably as seen in low AC participants using LOA 2.

VIII. General Discussion

When presented with an automation system that permitted choice of LOA in audiovisual temporal discrimination tasks, low AC participants generally favored the highest degree of selectable automation. Conversely, high AC participants opted to use the fully manual level more than the low AC group. These outcomes may provide preliminary evidence that AC is inversely proportional to preferences for LOA.

LOA 4 was used more in the dual-task than in the single-task conditions, indicating that task demands were offloaded to the machine when two concurrent sources of stimuli required attention. This result was likely driven by the visual-combined task in which both the low and high AC groups used LOA 4 more than in any of the other three tasks. The effect of load may have thus been more dependent on the modalities of the tasks, since participants continued to use LOA 2 heavily in the auditory-combined task but sought increased machine assistance for the more challenging visual-combined task.

Both groups demonstrated higher mean d' values in the auditory tasks. Also, higher LOAs were commonly chosen for the visual tasks, an indication that participants perceived the visual demands as more difficult. These results corroborate previous findings in which auditory signal detection performance has shown superiority over visual signal detection performance (e.g., [29], [30]). Auditory perception may be less vulnerable to distraction due to the inherent proximity of auditory stimulation to the ears (see [31]).

As evidenced by generally higher d' values in Experiment 2 compared to Experiment 1, performance was superior in fixed automation. This effect may have been due to the extra task demands of attending to and switching LOA options, not recognizing the need for increased assistance, and/or preference for a certain style of automation interaction despite it not being the most effective style.

AC effects upon d' were expected to manifest, especially in the dual-task conditions but were not observed. The perceptual discrimination tasks used here may have been insufficiently taxing, such that individual differences in AC did not emerge clearly. One possible reason is the use of two different modalities in the audiovisual combined task. With two separate stores, the visual and auditory modalities do not necessarily compete for the same limited processing resources [48]. Two concurrent intramodal tasks, however, may have been more demanding and consequently evocative of better performance in the high AC group. Future investigations could measure attentional allocation between extreme AC groups in a multitasking paradigm using intramodal stimuli.

High AC participants consistently exhibited a more conservative response bias than the low AC group, regardless of whether LOA was selectable or fixed. This effect was especially robust in the auditory modality, wherein they had a higher β than the low AC group, and their auditory β was higher than their visual β. High AC participants also appeared to be more influenced by the machine's neutral policy, perhaps due to greater ability to attend to and incorporate a decision aid.

Modality effects upon response bias were also influenced by having a choice of LOA. In Experiment 1, the visual and auditory β values were equivalent for low AC participants. However, in Experiment 2, their visual β significantly exceeded their auditory β. The low AC group, therefore, benefited from LOA choice, as modality differences emerged when the choice was removed. The high AC group, however, demonstrated the same modality effects for β in conditions of both adaptable and fixed LOAs. Across all tasks, the low AC group exhibited the most consistently neutral responses when they had a choice of LOA. It may be speculated that in terms of β, LOA choice may benefit those with low AC more, as they efficiently selected LOAs that enabled them to perceive targets and nontargets equally well.

IX. Practical Implications

The results of the present study indicate that an intermediary level of machine control was generally preferred across AC groups, possibly because it gave the operator default control, thereby encouraging task engagement, while also providing reliable confirmation to each of the operators' responses. A trend was observed for specific LOAs to complement the operators' unique abilities to regulate selective and divided attention as required. Although AC level may be generally inversely proportional to LOA preference, there may be other individual differences involved, such as desirability of control [49], which has been positively correlated with AC [50], [51]. Further, these operator characteristics may influence user acceptance of automation, which could significantly modulate LOA preference [52].

Adaptable automation may not always be the ideal solution, and preferences and needs may at times be incompatible. It may sometimes be more beneficial to let the system invoke automation, especially when the operator is too preoccupied to do so [8] or fails to assess the need for it [47]. Another important objective is to minimize demands associated with the control and supervision of adaptable automation [9] by developing proper usability and training for LOA management.

In designing complex human–machine systems, the inherently elevated salience of auditory stimuli should be recognized for its ability to both increase operator awareness and override the salience of competing visual stimuli. Therefore, whereas it is eminently possible that multimodal displays could aid the human operator (e.g., [1], [53]), these results highlight the potential for visual warnings to be neglected in favor of auditory ones. This prospect argues for the *a priori* consideration of psychophysical equivalency scaling in multimodal displays using visual, auditory, and indeed, tactile information presentation. Adaptable automation, however, could in part address intermodal differences by assisting in monitoring less salient stimuli.

The current study identifies the value of considering individual differences in operators and the consequent implications for automation invocation. While attentional behavior has demonstrated its effects upon operator-automation interaction, additional individual difference variables will almost certainly have important implications for automation technology in supporting a variety of human characteristics.

References

[1] C. E. Billings, *Aviation Automation: The Search for a Human-Centered Approach*. Mahwah, NJ, USA: Lawrence Erlbaum Assoc., 1997.

[2] P. A. Hancock and M. H. Chignell, "Adaptive control in human-machine systems," in *Human Factors Psychology*. P. A. Hancock, Ed. Amsterdam, The Netherlands: North Holland, 1989, pp. 305–345.

[3] T. B. Sheridan, *Human and Automation: System Design and Research Issues*. New York, NY, USA: Wiley, 2002.

[4] P. A. Hancock, "Imposing limits on autonomous systems," *Ergonomics*, vol. 60, no. 2, pp. 284–288, 2016.

[5] P. A. Hancock, M. H. Chignell, and A. Lowenthal, "KBAM: A prototype knowledge-based adaptive man-machine system," in *Erg. Int. 85: Proc. 9th Congr. of the Int. Erg. Assoc.*, I. D. Brown, R. Goldsmith, K. Coombes, and M. A. Sinclair, Eds. London: Taylor & Francis, 1985, pp. 103–105.

[6] W. B. Rouse, "Adaptive aiding for human-computer control," *Hum. Factors*, vol. 30, pp. 431–438, 1988.

[7] W. B. Cannon, "Organization for physiological homeostasis," *Physio. Rev.*, vol. 9, pp. 399–431, 1929.

[8] E. L. Wiener, "Human factors of advanced technology ("glass cockpit") transport aircraft," *NASA-Ames Research Center*, Moffett Field, CA, USA, Tech. Rep. 177528, Jun. 1989.

[9] D. D. Woods, "Decomposing automation: Apparent simplicity, real complexity," in *Automation and Human Performance: Theory and Applications*, R. Parasuraman and M. Mouloua, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Assoc., 1996, pp. 1–16.

[10] W. A. Olson and N. B. Sarter, "Automation management strategies: Pilot preferences and operational experiences," *Int. J. Av. Psych.*, vol. 10, no. 4, pp. 327–341, 2000.

[11] H. A. Ruff, S. Narayanan, and M. Draper, "Human interaction with levels of automation and decision aid fidelity in the supervisory control of multiple simulated teleoperated air vehicles," *Presence*, vol. 11, pp. 335–351, 2002.

[12] K. S. Tso, G. K. Tharp, A. T. Tai, M. H. Draper, G. L. Calhoun, and H. A. Ruff, "A human factors testbed for command and control of unmanned air vehicles" in *Proc. 22nd Digit. Avionics Syst. Conf.*, Indianapolis, IN, USA, Oct. 2003, pp. 8.C.1-81-12.

[13] L. J. Prinzel, F. G. Freeman, M. W. Scerbo, P. J. Mikula, and A. T. Pope, "Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component," *Hum. Factors*, vol. 45, no. 4, pp. 601–613, 2003.

[14] D. B. Kaber, M. C. Wright, L. J. Prinzel, and M. P. Claman, "Adaptive automation of human-machine system information-processing functions," *Hum. Factors*, vol. 47, no. 4, pp. 730–741, 2003.

[15] P. A. Hancock, G. M. Hancock, and J. S. Warm, "Individuation: The N = 1 revolution," *Theoretical Issues Erg. Sci.*, vol. 10, no. 5, pp. 481–488, 2009.

[16] M. R. Endsley, "From here to autonomy: Lessons learned from human–automation research," *Hum. Factors*, vol. 59, no. 1, pp. 5–27, Dec. 2016.

[17] F. Flemisch, M. Heesen, T. Hesse, A. Schieben, and J. Beller, "Toward a dynamic balance between human and automation: Authority, ability, responsibility, and control in shared and cooperative control situations," *Cognition, Technol. Work*, vol. 14, no. 1, pp. 3–18, Mar. 2012.

[18] D. E. Egan, "Individual differences in human-computer interaction," in *Handbook of Human-Computer Interaction*, M. Helander, Ed. Amsterdam, The Netherlands: North Holland, 1988.

[19] G. Matthews, D. R. Davies, S. J. Westerman, and R. B. Stammers, *Human Performance: Cognition, Stress, and Individual Differences*. Philadelphia, PA, USA: Taylor & Francis, 2000.

[20] J. L. Szalma, "Individual differences in performance, workload, and stress in sustained attention: Optimism and pessimism," *Personality Indiv. Differences*, vol. 47, pp. 444–451, 2009.

[21] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. DeVisser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Hum. Factors*, vol. 53, pp. 517–527, 2011.

[22] K. A. Hoff and M. Bashir, "Trust in automation: Integrating empirical evidence on factors that influence trust," *Hum. Factors*, vol. 57, pp. 407–434, 2015.

[23] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Trans. Syst., Man, Cybern., Part A*, vol. 30, no. 3, pp. 286–297, May 2000.

[24] D. Derryberry and M. A. Reed, "Anxiety-related attentional biases and their regulation by attentional control," *J. Abnormal Psych.*, vol. 111, no. 2, pp. 225–236, 2002.

[25] M. W. Eysenck, N. Derakshan, R. Santos, and M. G. Calvo, "Anxiety and cognitive performance: Attentional control theory," *Emotion*, vol. 7, no. 2, pp. 336–353, 2007.

[26] T. B. Sheridan, "Task allocation and supervisory control," in *Handbook of Human-Computer Interaction*, M. Helander, Ed. Amsterdam, The Netherlands: North Holland, 1988.

[27] D. Navon and D. Gopher, "On the economy of the human processing systems," *Psychol. Rev.*, vol. 86, pp. 254–255, 1979.

[28] P. A. Hancock, "The effects of automation invocation procedure and dynamic display relocation on performance in a multi-task environment," *IEEE Trans. Syst., Man, Cybern., Part A*, vol. 37, no. 1, pp. 47–57, Jan. 2007.

[29] J. L. Szalma *et al.*, "Effects of sensory modality and task duration on performance, workload, and stress in sustained attention," *Hum. Factors*, vol. 46, pp. 219–233, 2004.

[30] J. S. Warm and H. J. Jerison, "The psychophysics of vigilance," in *Sustained Attention in Human Performance*, J. S. Warm, Ed. New York, NY, USA: Wiley, 1984, pp. 15–59.

[31] J. L. Hatfield and M. Loeb, "Sense mode and coupling in a vigilance task," *Perception Psychophys.*, vol. 4, pp. 29–36, 1968.

[32] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.

[33] B. G. Tabachnick, L. S. Fidell, and S. J. Osterlind, *Using Multivariate Statistics*. Boston, MA, USA: Allyn & Bacon, 2001.

[34] M. L. Dittmar, J. S. Warm, W. N. Dember, and D. F. Ricks, "Sex differences in vigilance performance and perceived workload," *J. Gen. Psych.*, vol. 120, pp. 309–322, 1993.

[35] S. Droit-Volet, S. Tourret, and J. Wearden, "Perception of the duration of auditory and visual stimuli in children and adults," *Quart. J. Exp. Psych.*, vol. 57, no. 5, pp. 797–818, 2004.

[36] W. M. Dember and T. S. Warm, *The Psychology of Perception*, 2nd ed. New York, NY, USA: Wiley, 1979.

[37] G. Gescheider, *Psychophysics: The Fundamentals*, 3rd ed. Mahwah, NJ, USA: Lawrence Erlbaum Assoc., 1997, p. ix.

[38] S. R. Dixon and C. D. Wickens, "Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload," *Hum. Factors*, vol. 48, no. 3, pp. 474–486, 2006.

[39] J. J. Goh and D. A. Wiegmann, "Effects of automation failures during a visual inspection task on reliance and visual skill acquisition: A comparison between direct and indirect cueing," *Human Factors Division Inst. of Aviation*, Savoy, IL, USA, Tech. Rep. HFD-06-06, Oct. 2006.

[40] S. M. Galster, R. S. Bolia, and R. Parasuraman, "The application of a qualitative model of human interaction with automation: Effects of unreliable automation on performance," Air Force Research Lab, Wright-Patterson AFB, OH, USA, Tech. Rep. ADA422127, Oct. 2003.

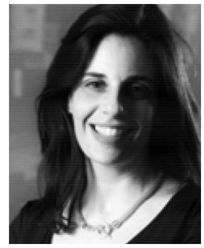
[41] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York, NY, USA: Wiley, 1966.

[42] N. A. Macmillian and C. D. Creelman, *Detection Theory: A User's Guide*, 2nd ed. Mahwah, NJ, USA: Lawrence Erlbaum Assoc., 2005.

[43] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” *Adv. Psych.*, vol. 52, pp. 139–183, 1988.
[44] J. C. Byers, A. C. Bittner, and S. G. Hill, “Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary?” in *Adv. Ind. Erg. & Safety*, A. Mittal, Ed. London: Taylor & Francis, 1989.
[45] T. E. Nygren, “Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload,” *Hum. Factors*, vol. 33, no. 1, pp. 17–33, 1991.
[46] M. K. Tulga and T. B. Sheridan, “Dynamic decisions and workload in multitask supervisory control,” *IEEE Trans. Man Cybern.*, vol. 10, no. 5, pp. 217–232, May 1980.
[47] N. M. Morrison and W. B. Rouse, “Adaptive aiding for human-computer control: Experimental studies of dynamic task allocation,” Naval Air Warfare Center, Warminster, PA, USA, Tech. Rep. AAMRL-TR-86-005, Jan. 1986.
[48] C. D. Wickens, “Multiple resources and performance prediction,” *Theoretical Issues Erg. Sci.*, vol. 3, no. 2, pp. 150–177, 2002.
[49] J.M. Burger and H.M. Cooper, “The desirability of control,” *Motivation Emotion*, vol. 3, pp. 381–393, 1979.
[50] J. E. Thropp, T. Oron-Gilad, J. L. Szalma, and P. A. Hancock, “Incorporating individual differences into the adaptive automation paradigm,” presented at the 113th APA Convention, Div. 21, Washington, DC, USA, Aug. 18–21, 2005.
[51] J. E. Thropp, T. Oron-Gilad, J. L. Szalma, and P. A. Hancock, “Low and high levels of attentional control in adaptive automation,” presented at the 50th Hum. Factors Erg. Soc. Annu. Mtg., San Francisco, CA, USA, Oct. 16–20, 2006.
[52] M. Jipp, “Levels of automation: Effects of individual differences on wheelchair control performance and user acceptance,” *Theoretical Issues Erg. Sci.*, vol. 15, no. 5, pp. 479–504, 2013.
[53] N. B. Sarter and D. D. Woods, “Pilot interaction with cockpit automation II: An experimental study of pilots’ model and awareness of the flight management system,” *Int. J. Av. Psych.*, vol. 4, pp. 1–28, 1994.

Jennifer E. Thropp received the B.S. degree in psychology in 2001, the M.S. degree in modeling and simulation in 2004, and the Ph.D. degree in applied/experimental psychology from the University of Central Florida, Orlando, FL, USA.

She is an Assistant Professor with the School of Graduate Studies, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA. Her research interests include human perceptual, cognitive, and attentional performance in high-altitude conditions.



Tal Oron-Gilad (M’10) received the B.Sc. and M.Sc. degrees in industrial engineering from the Technion in 1992 and 1997, respectively, and the Ph.D. degree in human factors engineering in 2003 from the Ben-Gurion University, Beersheba, Israel.

She is a Professor and the Chair with the Department of Industrial Engineering and Management, Ben-Gurion University, Beersheba, Israel. Her research interests include human factors engineering and decision support systems.

She has been funded by the US Army Research Lab, General Motors, the Israeli Aerospace Industry, the Israeli Road Safety Authority, Israeli Ministry of Science, and EU projects (SOCRATES).

James L. Szalma received the B.S. degree in chemistry in 1990 from the University of Michigan, MI, USA, and the M.A. and Ph.D. degrees in applied experimental/human factors psychology in 1997 and 1999, respectively, from the University of Cincinnati, Cincinnati, OH, USA.

He is an Associate Professor with the Psychology Department, University of Central Florida, Orlando, FL, USA and the Director with the Performance Research Laboratory. His research interests include signal/threat detection, training for threat detection, fuzzy signal detection theory, and how the characteristics of tasks and operators interact to influence performance.

Peter A. Hancock (A’05–M’06–F’17) received the B.Ed. in anatomy and physiology in 1976, the M.Sc. in human biology in 1978, and the D.Sc. in human-machine systems in 2001, all from Loughborough University, Loughborough, England, and the Ph.D. in human performance in 1983 from the University of Illinois, Champaign, IL, USA.

He is a Provost Distinguished Research Professor with the Department of Psychology and with the Institute for Simulation and Training, University of Central Florida, Orlando, FL, USA. He has authored over 900 refereed scientific articles and publications. His research interests include human interaction with autonomous systems.