Objective: In this work, we expand on the theory of adaptive aiding by measuring the effectiveness of coadaptive aiding, wherein we explicitly allow for both system and user to adapt to each other.

Background: Adaptive aiding driven by psychophysiological monitoring has been demonstrated to be a highly effective means of controlling task allocation and system functioning. Psychophysiological monitoring is uniquely well suited for coadaptation, as malleable brain activity may be used as a continuous input to the adaptive system.

Method: To establish the efficacy of the coadaptive system, physiological activation of adaptation was directly compared with manual activation or no activation of the same automation and cuing systems. We used interface adaptations and automation that are plausible for real-world operations, presented in the context of a multi–remotely piloted aircraft control simulation. Each participant completed 3 days of testing during 1 week. Performance was assessed via proportion of targets successfully engaged.

Results: In the first 2 days of testing, there were no significant differences in performance between the conditions. However, in the third session, physiological adaptation produced the highest performance.

Conclusion: By extending the data collection across multiple days, we offered enough time and repeated experience for user adaptation as well as online system adaptation, hence demonstrating coadaptation.

Application: The results of this work may be employed to implement more effective adaptive workstations in a variety of work domains.

Keywords: neuroergonomics, mental workload, function allocation, human/computer interaction

INTRODUCTION

Adaptive aiding and adaptive automation have been studied and discussed for quite some time. The idea of reallocating tasks between human and machine (Rouse, 1976) to produce maximum overall system performance has an intuitive appeal that has led to great interest through the years. There are many possible methods of determining task allocation in an adaptive system, but one that has shown particular promise is based on psychophysiological monitoring (Byrne & Parasuraman, 1996; Parasuraman & Wilson, 2008; Pope, Bogart, & Bartolome, 1995; Scerbo, 2007) and real-time modification (Wilson & Russell, 2003, 2007). The present study was intended to build on this work in two key areas: increasing the realism of the adaptive aiding used and extending data collection across multiple sessions and days. This extension across days allows time for the operator to adapt to the system as well as adaption of the system to the moment-by-moment capabilities of the operator, demonstrating coadaptation. To strongly test for improvements with physiologically triggered aiding, physiologically triggered adaptive aiding was compared with manually triggered adaptable aiding, on the basis of work by Bailey, Scerbo, Freeman, Mikulka, and Scott (2006).

Wilson and Russell (2007) presented a particularly effective demonstration of the potential of real-time, physiologically activated adaptive aiding. They used a complex simulation in which participants operated multiple remotely piloted aircraft (RPA) on a ground attack mission while electroencephalogram (EEG), electrocardiogram (ECG), and electrooculogram (EOG) data (hereafter referred to collectively as physiological data) were monitored. The physiological data were classified in real time by trained artificial neural networks. The neural networks continuously classified operator workload and initiated aiding when high workload was detected. The aiding consisted of reducing the airspeed (by 50%) of those vehicles that were approaching a
target. They demonstrated that this aiding resulted in significantly improved operator performance when activation was based on the physiological workload rather than random activation. This study provided a strong demonstration of the potential positive impact of real-time activation of aiding based on physiologically assessed workload.

Reducing the airspeed by half and consequently reducing the rate at which events had to be processed is clearly a highly effective means of mitigating task demands on the operator. However, in many application settings, this type of task slowing may not be a realistic option. Extensively developing effective and realistic demand reduction techniques was not, in fact, one of Wilson and Russell’s (2007) objectives, as the primary goal was to demonstrate the potential of such adaptive aiding systems. To further advance this area of work, it is necessary to demonstrate the effectiveness of more realistic aiding techniques appropriate to multi-RPA operation. To address this need, the present study used an adaptive system that implemented a suite of mitigation procedures corresponding to Stages 2 and 3 in the hierarchy proposed by Fuchs, Berka, Levendowski, and Juhnke (2006): directing attention via cuing and enhancing the salience of critical events via decluttering. In addition, limited automation was invoked that paired vehicles and targets, subject to operator override. This work builds on a similar study performed by Parasuraman, Cosenzo, and De Visser (2009) by testing attentional aiding in addition to automation.

A key potential benefit offered by physiological activation is that the system should have little or no workload associated with managing activation (Byrne & Parasuraman, 1996), unlike manually activated aiding. However, no physiologically based workload monitoring system has yet achieved perfect detection of high workload, at least partly because of noise or artifacts (Smith, Gevins, Brown, Karnik, & Du, 2001). Consequently, such systems exhibit errors and activate aiding when it is unnecessary or possibly detrimental to performance. Although the operator could make similar errors in manually activating the aiding system, a critical question is whether overall performance achievable with an imperfect physiological activation system exceeds that possible with manual activation. Undoubtedly, the answer depends on the accuracy of physiologically driven activations as well as the workload involved in manual management of activation. Bailey et al. (2006) demonstrated modest performance improvements for physiologically activated automation as compared with manually activated; this study builds on that work by allowing significantly greater time for participants to adapt to the physiological activation system.

There is a close relation between physiologically activated adaptive aiding and brain–computer interfaces (BCIs). BCI here refers to the use of brain signals to directly control systems, such as the classic example of communicating via selection of letters based on analysis or classification of EEG signals (Farwell & Donchin, 1988). Physiologically activated adaptive aiding is, in a sense, a special case of BCI wherein the purpose is not direct control but rather monitoring and providing aiding to operators to enable them to work more effectively (often referred to as passive BCI; e.g., Zander, Kothe, Jatzev, & Gaertner, 2010). Researchers investigating BCI have noted that brain signals used for BCI applications generally evolve with feedback in a manner that improves system performance. Recognizing that brain signals are adapting in BCI paradigms suggests that similar ongoing adaptation of analysis techniques or classifiers (e.g., Wolpaw, McFarland, Neat, & Forneris, 1991) is a promising approach.

However, BCI work demonstrating the value of recognizing that both the operator and physiological classification systems are adaptive controllers has been accomplished only recently (Huang, Erdogmus, Pavel, Mathan, & Hild, 2011; McFarland, Sarnacki, & Wolpaw, 2011). In this line of work, operators who were trained to use a BCI system across a period of days achieved greater accuracy with practice. Analysis revealed that operators adapted their own EEG signals to more precisely trigger the system. This mutually adaptive relationship in BCI (as proposed by Wolpaw, Birbaumer, McFarland,
Pfurtscheller, & Vaughan, 2002) may then be extended to application in a physiologically activated adaptive interface. In an adaptive interface, easily visible interface changes provide the necessary feedback to operators who may then adapt brain activity to improve their management of the interface. Conversely, the physiological activation system may be adapted to the operator by retraining classifiers to capitalize on changes in user brain activity. This study retrained classifiers for each participant and each testing day to adapt to such changes. We term this application and overall process coadaptation; hence in the present study, we tested an implementation of a coadaptive interface.

The extension of data collection to multiple days was motivated by previous results in which BCI-adapted EEG was shown to induce neuroplasticity (Ros, Munneke, Ruge, Gruzelier, & Rothwell, 2010) and by separate work that has demonstrated that neural activity during sleep is a key factor in plasticity and the consolidation of learned items (Steriade, 2001). The 1st and 2nd days of collection in the current study cover a normal sleep-wake cycle and were chosen as most likely to support adaptive changes in the operator’s neural signals. The 3rd day of data collection was spaced to 1 week after the 1st day; this spacing was based on learning research that has demonstrated that expanding the interstudy interval can lead to more stable retention (Hser & Wickens, 1989), while best accommodating participant availability.

In the present study, therefore, we sought to compare performance in a realistic task with a physiologically activated adaptive aiding system with the same system when triggered by a simple manual key press. The study sessions were conducted across 3 days to facilitate neuroplastic changes associated with normal sleep-wake cycles and expanding interstudy intervals. By thus supporting user adaptation to an already-adaptive system, the study ultimately facilitated coadaptation.

**METHOD**

**Participants**

Participants were 10 naive persons either currently employed at Wright-Patterson Air Force Base or students at Wright State University in Dayton, Ohio, who volunteered to participate in the study. (An additional 6 participants, 5 male and 1 female, were enrolled in the study. The total duration of the study, including training and testing, typically extended across several months. Because of the duration, these participants left prior to completion, most commonly because they moved or graduated.) Employees of Wright-Patterson Air Force Base received their normal duty pay, and Wright State students were paid $15 per hour for their participation. Participants completed comprehensive written informed consent prior to the start of the experiment. All reported normal or corrected-to-normal vision with no color blindness. Of the 10 participants, 2 were female and 8 were male, all between the ages of 19 and 26 years (mean of 22 years). All study procedures were reviewed and approved by the Air Force Research Laboratory Institutional Review Board.

**Apparatus and Stimuli**

The multi-RPA operation task was a PC-based supervisory control simulation of a mission involving “suppression of enemy air defense” (Schmidt, Wilson, Funke, Davis, & Caldwell, 2010) using keyboard and mouse controls. Participants monitored the progress of 8 or 16 generic RPAs on two abutted 51-cm (diagonal) computer screens as they flew a preplanned mission. When the RPAs came within radar range of a target, simulated radar images of the target area were automatically acquired, and operators could then mouse-click a map icon to download and view these images. Each image contained zero to eight targets drawn from three visually distinct types as well as 25 to 30 non-target distractors. Each target type was to be engaged with a specific weapon type, generically termed small, medium, and large. Each RPA carried a limited number of two of the three weapon types. After visually searching the images, participants were required to select each weapon with the mouse and click the image to mark appropriate weapon–target pairings before the RPA reached the minimum weapons release distance.

These tasks were performed for each of the RPAs as targets came in range. If the targets were not selected and/or the weapons release command (mouse click on a confirmation box)
was not given in time, the weapons from that RPA could not be released, thereby reducing the number of targets successfully engaged. Likewise, misidentifying a target and assigning an inappropriate weapon type was not counted as success. Participants could use the mouse to designate waypoints and direct RPAs away from preplanned routes but were not allowed to double back to reengage targets. Successfully engaging all targets required rerouting RPAs as stores were expended. This task was intended to broadly represent future operator control tasks and tap a wide range of cognitive skills, such as working memory, visual search, object recognition, task switching, and flexible management of conflicting priorities.

Adaptive aiding and automation were implemented in this task via several methods. The aiding methods were intended to cue attention to time-critical tasks. Because the aircraft could not double back, a simple time-to-contact calculation enabled prioritization of those targets that most urgently needed operator attention. The top three target–RPA pairs were then color coded with red, yellow, and green transparent circles. All other RPAs and targets had contrast lowered (“fog layer”) to render them less salient (Figure 1). In addition to these aiding techniques, the interface supported partial task automation. Instead of the operator’s manually assigning an RPA to a target, the automation linked each target to the closest RPA. This linking was made without regard to the weapon type required or stores on that RPA and thus could link RPAs that were incapable of engaging a target. The operator could see the stores on each RPA and override incorrect links. Last, when a target came into range, the automation displayed the appropriate simulated radar image rather than requiring the operator to request it. All of the aiding techniques and automation were activated and deactivated together, either on operator command or on the basis of

Figure 1. An example of the multi–remotely piloted aircraft (RPA) control task. The wedge-shaped symbols at the bottom represent RPAs that the operator is controlling; the missile-shaped symbols represent air defense sites that need to be engaged. Aircraft proceeded from south to north at a fixed rate, requiring approximately 20 min to traverse the target area. The color cuing is active in this example and is visible as same-color circles around a target–RPA pair.
physiological workload classification. Operator command (the manual activation condition) consisted of pressing the space bar on the keyboard. Physiological activation was triggered on the basis of output of a physiological workload classifier that replicated Wilson and Russell (2007).

A concern with the workload classification–based aiding is that effective aiding should reduce workload, leading to deactivation of the aiding. This effect could then lead to a high frequency of activation and deactivation as workload is modulated by the aiding. To reduce this occurrence, any change in aiding (either manually activated or physiologically activated) triggered a time-out of 15 s before another such change was allowed. Aiding remained on if reactivation was commanded during this 15-s window.

Physiological data recording and analysis replicated Wilson and Russell (2007). Briefly, we included EEG from five channels, at F7, Fz, Pz, T5, and O2, positioned according to the International 10-20 electrode system (Jasper, 1958), using an Electro-Cap (Electro-Cap International, Inc., Eaton, OH). Reference and ground electrodes were positioned on the mastoid processes, with impedances verified below 5 kiloohms. Horizontal and vertical EOG (HEOG and VEOG, respectively) and ECG were also recorded with the use of standard Ag/AgCl electrodes. A Cleveland Medical Devices, Inc. (now Great Lakes NeuroTechnologies, Cleveland, OH), BioRadio 110 telemetry unit was used to acquire these data channels with a sampling rate of 200 Hz (12-bit resolution, band-pass filtered between 0.5 and 52.4 Hz). Corrections for eye movement and blinks were made with the use of an online implementation of an adaptive filter with HEOG and VEOG used as reference noise channels (He, Wilson, & Russell, 2004; He, Wilson, Russell, & Gerschutz, 2007). Interbeat interval (IBI), calculated across a 10-s window, was derived from the ECG channel with the use of an online algorithm (Hamilton & Tompkins, 1986; Pan & Tompkins, 1985). Similarly, blink rate, calculated across a 30-s window, was derived with the use of an online algorithm developed by Kong and Wilson (1998).

The EEG data were filtered into separate bandlimited channels with elliptical infinite impulse response filter banks. The passbands for each channel were consistent with the five traditional bands of EEG: delta (0.5–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz), and gamma (31–42 Hz). For real-time classification, the data were segmented into 10-s windows with a 9-s overlap. Log power of the five bands from the five sites was used in addition to IBI and blink rate, resulting in 37 features as inputs for workload classification. Workload classification was accomplished via a feed-forward artificial neural network trained via back-propagation with two output nodes corresponding to low and high workload, replicating the procedure used by Wilson and Russell (2003). Immediately following each trial, participants completed the computerized NASA–Task Load Index (NASA-TLX; Hart & Staveland, 1988).

Procedure

Participants were extensively trained on the task prior to the experimental sessions to reduce training effects during the three data collection days. Each participant completed a minimum of 2 hr of training per day for 5 days. During the course of these days of training, participants were gradually introduced to all of the task features. The aiding system was introduced on approximately the 4th day of training; the 4th and 5th days included practice manually controlling the aiding system. Participants were instructed to maximize their performance by using aiding as much or as little as they liked when it was available. Training was continued until a performance criterion (80% of targets engaged without any aiding while controlling 16 RPAs) was met.

Testing days for each participant followed a fixed schedule: Day 2 was always right after Day 1, with Day 3 one week after Day 1 (6 days after Day 2). This schedule was chosen to maximize the potential for operator adaptation in using the adaptive aiding system, space the learning sessions across an expanding interval, and maximize participant availability. On each testing day, after fitting of all electrodes, the participants completed one 5-min task run to warm up. A subsequent resting baseline included 2 min of participants’ sitting with their eyes
open, blinking normally, and 2 min of sitting with eyes closed. Setting up the physiologically activated aiding then required two classifier training trials, one 10-min trial each of low difficulty (8 RPAs) and high (16 RPAs), with no aiding. To train the classifier, 20 minutes of data were sufficient; however, significantly fewer data (10 min total) have proven sufficient in other similar studies (Christensen, Estepp, Wilson, & Russell, 2012). These classifier training trials were repeated on each new day of testing to adapt the classifier to changes that may have taken place in operator brain signals. To avoid order or carryover effects, we randomized the order of trials for each participant and day. However, the classifier training trials had to precede any run including physiologically activated aiding.

With this constraint and three experimental trials (one each of no aiding, physiologically activated aiding, and manual activated aiding), there are 72 possible randomizations. We selected 30 that resulted in each participant’s completion of the three aiding conditions in a unique order each day, with the physiologically activated condition appearing once in each possible position. The experimental trials were 20 min in duration, one each with no aiding, manually activated aiding, and physiologically activated aiding. Each experimental trial included a low-difficulty middle segment of approximately 5 min during which only 8 of the 16 RPAs had targets to engage; the remainder of the trial was high difficulty, with all 16 RPAs engaging targets. A total of 120 targets were presented in these high-difficulty segments. The embedded low-difficulty segment enabled verification that the physiologically activated aiding did not simply remain on for the entirety of a run. Participants completed the NASA-TLX immediately following each trial. In total, the testing session took approximately 5 hr each day. Performance data will be reported for the high-difficulty (16 RPAs controlled) portions only, as low-difficulty portions are at ceiling.

RESULTS

Performance was calculated as the average proportion of targets successfully engaged in the high-difficulty segments of a trial (out of 120 high-difficulty targets total per trial). For statistical analysis, these proportions were corrected via the arcsine transform (Freeman & Tukey, 1950); all data reported in figures are raw proportions to aid in interpretation. A 3 (days) × 3 (no aiding, physiologically activated, manually activated) repeated-measures ANOVA was conducted with the Huynh-Feldt correction for any violations of sphericity. The main effects of aiding activation type and day of testing were not significant, \( p > .1 \). However, there was a significant interaction between the two factors, \( F(3.06, 27.6) = 5.37, p = .005 \), partial \( \eta^2 = .374 \).

It is evident from Figure 2 that the aiding groups did not differ on the first 2 days, but physiological activation improved on the 3rd day. We probed this interaction with individual ANOVAs comparing the three aiding conditions separately for each day; for the first 2 days, these tests were not significant, \( p > .1 \). On the 3rd day, this test was significant, \( F(2, 18) = 5.83, p < .01 \), partial \( \eta^2 = .421 \). Post hoc comparisons using the Tukey HSD test indicated that the mean for the physiological activation condition (\( M = .90, SD = .05 \)) was significantly greater than the means for the manual activation (\( M = .84, SD = .06 \)) and no-aiding conditions (\( M = .82, SD = .07 \)). The latter

![Figure 2. Performance results expressed as the mean proportion of total targets successfully engaged in the high-workload portions of a trial. Operator = operator control of aiding activation; Physio = physiological activation of aiding. Error bars are standard errors.](https://example.com/figure2.png)
two conditions were not significantly different from each other. The magnitude of the difference between physiological activation and manual activation on the 3rd day was modest at 6%; however, the effect size (Cohen’s $d$) for this comparison was large at 1.17.

The NASA-TLX subjective workload data were analyzed similarly to the performance data reported earlier. The overall scores were calculated via the unweighted procedure (Nygren, 1991). The task was indeed challenging; the mean overall score for high-difficulty conditions was 61 as opposed to 24 for the low-difficulty conditions. There was a great deal of apparently random variance in scores within participants across days, even at constant performance; to ameliorate this variance, scores were normalized via $z$ scoring within participant and day as recommended by Stevens (1971). Means and standard deviations for $z$ scoring were drawn from all eight composite scores (within participant and day): two from the calibration conditions and two from each of the three aiding conditions (one for the low-difficulty segment and one for the high-difficulty segment). A possible consequence of this normalization is a reduction in between-day differences if those differences are scalar; as we did not observe a main effect of days on performance, this potential reduction was judged acceptable. Even after normalizing, 1 participant produced outlier NASA-TLX scores that were approximately four standard deviations from the mean; this participant was excluded from all subsequent analyses of these data.

A $3 \times 3$ (three aiding conditions and three days) repeated-measures ANOVA revealed a significant main effect of day, $F(1.8, 14.5) = 13.01, p = .001$, partial $\eta^2 = .619$, and a significant interaction, $F(2.8, 22.3) = 3.09, p = .035$, partial $\eta^2 = .296$ (Figure 3). As with the performance data, individual ANOVAs compared the aiding conditions separately for each of the three days. These revealed no significant differences on Day 1 or Day 2 but a significant main effect on Day 3, $F(2, 16) = 12.08, p = .001$, partial $\eta^2 = .602$. Post hoc comparisons using the Tukey HSD test indicated that all three aiding conditions are significantly different from each other, with no aiding producing significantly higher workload ($M = 1.2, SD = 0.2$) than the operator-activated condition ($M = .94, SD = .2$), which in turn was higher than the physiological activation condition ($M = .69, SD = .2$).

As the activation of the aiding was not controlled across conditions, it was necessary to test whether the difference in performance may be simply explained by a difference in the proportion of time the aiding was active. A $2 \times 3$ (Aiding Activation Type × Day) repeated-measures ANOVA was run on the proportion of the high-workload segments of a trial during which the aiding was active. This ANOVA did not result in any significant effects; the closest to significance was the Aiding × Day interaction, $F(1, 9) = 2.50, p > .1$. Given the significant differences between physiological activation and the other conditions on the 3rd day reported earlier, the difference between operator activation and physiological activation was checked on Day 3 via a simple paired $t$ test. This result was likewise not significant, $t(9) = 1.01, p > .3$ (two tailed). For Day 3, operators activated the aiding for an average of 49% of the run, whereas the physiological activation resulted in aiding for 61% of the high-workload portions of a trial. This same analysis also indicated that the physiological activation system discriminated between the high- and low-workload segments; it was active for an average of 23% of the time during the low-workload segments. Given the
variable nature of participant workload during a run (because of differences in individual strategy, RPA routing, weapon stores usage, etc.), it is difficult to assign a “correct” classification percentage for the physiological activation system. It is nonetheless intuitive that aiding should be more active in segments of increased workload.

The 15-s time-out on activation and deactivation of aiding reduced but did not eliminate more frequent cycling of aiding in the physiologically activated condition. This finding was probed with a $2 \times 3$ (Aiding Activation Type $\times$ Day) repeated-measures ANOVA conducted on the number of activations. There was a significant main effect of activation type, $F(1, 9) = 35.09, p < .01$. Across days, aiding was triggered on average of five times per run in the manually activated aiding condition and 15 times in the physiologically activated condition. There was no significant effect of testing day or interaction effect ($p > .1$). Aiding cycled more frequently in the physiologically activated condition but was not on for a significantly greater length of time.

**DISCUSSION**

In this work, we set out to test the efficacy of simultaneously adapting an interface to the user via physiological monitoring while facilitating the possibility of neuroplastic changes associated with a normal sleep-wake cycle and interstudy intervals. The significant interaction between type of aiding and days of testing reveals that this approach indeed resulted in improved performance beyond that achieved with a simple manually controlled adaptable interface, which is generally consistent with Bailey et al. (2006). This improvement is perhaps the key criterion for a successful physiological aiding system: Despite imperfect workload classification, it is still superior to manual control.

Performance was not significantly different across the 3 days in the no-aiding and the manually activated aiding conditions. This finding validates that task training did indeed approach asymptote; if participants were still learning the task, we would have expected a simple main effect of days, regardless of aiding condition. Similarly, the training with the manually activated aiding was effective; if they were still learning to use the aiding system, we would expect an interaction between aiding condition and days, with significantly better performance in both of the aided conditions relative to the unaided condition. There were no significant differences between manually activated aiding and no aiding in performance; this finding is evidence that learning or adaptation is confined to physiologically activated aiding, consistent with our hypothesis regarding coadaptive aiding.

Analysis of the NASA-TLX workload scores revealed that subjective workload mirrors performance to some degree. The same interaction between aiding and day was observed, although post hoc testing revealed that all three aiding conditions were different from each other on Day 3, whereas in the performance analyses, operator activation and no aiding were not significantly different from each other on Day 3. The pattern of differences in NASA-TLX on Day 3 matches expectations, with no aiding producing the highest workload, operator activation the next highest, and physiological the lowest. This finding confirms that on Day 3, the physiological activation produced both the highest performance and the lowest subjective workload. The manually activated aiding is not completely ineffective; subjective workload was lower on Day 3 than that observed with no aiding, even though there was no significant difference in performance.

The present work was not designed to fully elaborate the nature of adaptive changes occurring within the operator. However, we may reasonably infer that the process may involve strategic changes in their approach to the task, changes in physiological signals that reduce noise (McFarland et al., 2011), and perhaps the development of conscious control of signals to activate aiding based on what may be considered a biofeedback paradigm. Strategic changes would have to be uniquely effective with physiologically activated aiding to produce the observed pattern of results and were not reported by participants in poststudy debriefing. On the other hand, greater facility with the physiological activation system as a result of adaptive changes
in their own physiological signals is consistent with the existing literature on neuroplasticity and adaptive BCI. It is also not possible to determine to what degree coadaptation requires conscious effort or intention to adapt; there is a large body of literature on implicit learning (Reber, 1989) that suggests conscious effort may not be required. As all such adaptive changes on the part of the operator presumably have a neural substrate, it may not be possible to separate the exact causes or sources of coadaptation; that is, a reduction in noise in physiological workload classification could be an ancillary effect of neural changes associated with strategic change.

Although the proportion of time in which aiding was active was equivalent between the manually activated and physiologically activated aiding conditions, this equivalency was achieved with more on-off cycling in the physiologically activated condition. This finding is consistent with the expectation that effective aiding should reduce workload, leading to deactivation of the aiding, although simple errors in workload classification may also contribute. These factors result in more frequent transitions both from low to high workload and from high to low workload; the performance effects of these transitions may have canceled each other out (Matthews, 1986) or produced a net decline in performance (Krulewitz, Warm, & Wohl, 1975) relative to the less frequent transitions in the manually activated aiding condition. It is therefore possible that the observed improvement in performance with physiologically activated aiding would have been increased with less frequent transitions, perhaps achieved via a longer time-out between such changes. The management of workload transitions in an adaptive aiding context will require careful attention in future work.

The overall improvement in performance obtained with this realistic combination of automation and task cuing was nonetheless relatively modest. At best (Day 3, with physiological activation as compared with no aiding), the proportion of targets engaged improved from .81 to .90, or 9%. However, this improvement is similar to the 12% improvement observed in Bailey et al. (2006) and the 15% improvement observed in Prinzel, Freeman, Scerbo, Mikulka, and Pope (2000), both with more artificial tasks. Although this result may be coincidental, it is possible that the decreased workload associated with not having to manually manage an adaptive interface in a complex task should result in these levels of modest performance improvements.

It is also of note that the manually controlled aiding resulted in performance no better than that observed with no aiding at all. By chance, this study may have happened upon a combination of aiding effectiveness, workload associated with managing the aiding, and task difficulty that resulted in the additional management workload’s canceling out any benefits from having the aiding. If that result is true, then the physiologically activated condition reveals the underlying effectiveness of the aiding system, as it does not burden the operator with additional workload to manage activation. We may still then conclude that the aiding used in this study is of relatively marginal benefit, which points at a fundamental limitation in this line of research: Many or most system adaptations and automation that improve performance overall in a task are interface improvements that should be used at all times. The promise of adaptive interfaces is predicated on the assumption that there is some cost associated with constant usage, such that performance is less than optimal. For the experimenter who wishes to study adaptive aiding in realistic tasks, this assumption creates a challenging problem in expert system and interface design: develop aiding or automation that is useful under condition of high load but detrimental under low load.

In this study, this condition was met by the use of imperfect, simple, rule-based automation; a nonoverloaded operator could outperform the automation and thus should not use it. As has been discussed (Fuchs et al., 2006), in tasks such as RPA operations, a considerable amount of state information must be available for automation to be effective. In this study, the state of the environment is known absolutely, and a reasonable approximation of ideal user behavior is possible because of the constrained nature of the task. To find real-world application, adaptive interface techniques must either be constrained to a very narrow window of user
behavior (i.e., avoiding controlled flight into terrain as in automatic ground collision avoidance) or use artificial intelligence that is considerably more sophisticated than systems widely available now. On the basis of Wilson and Russell (2007), much better performance improvements are achievable with physiological activation of highly effective aiding techniques; advancement in this area will be required for operationally effective adaptive interfaces.

Outside operational applications, the challenge of implementing effective and appropriate aiding may be avoidable. In training or team-based applications (e.g., Elkins et al., 2009; Espevik, Helge Johnsen, Eid, & Thayer, 2006; Stevens, Galloway, Berka, & Sprang, 2009), it may be sufficient to provide additional state information via physiological workload monitoring, thus enabling either a human instructor or human teammates to more effectively adapt their own behavior or adapt the team composition more appropriately to the task (e.g., Woolley et al., 2007).

In summary, the present work demonstrated the effectiveness of realistic coadaptive aiding in a simulated multi-RPA control task. Over time, the users adapted their interaction with physiologically activated aiding, while that interface adapted to them, hence demonstrating coadaptive aiding.

ACKNOWLEDGMENTS

The authors wish to thank Iris Davis, Justin Cheuvront, and Margaret Bowers for their considerable efforts in implementing and running this study. The authors would also like to thank Glenn Wilson for providing insightful comments on the design and analysis and Joel Warm for many comments that have improved this manuscript. Funding for this work was provided by the Air Force Office of Scientific Research.

KEY POINTS

• Performance with physiologically activated adaptive aiding on the 3rd day of testing exceeded performance with manually activated aiding or no aiding.

• There was no significant effect of activation type on the first 2 days.

• Manually controlled aiding is a key comparison for alternative control mechanisms.

• Effective adaptation or mitigation techniques are difficult to design and implement in realistic tasks.

REFERENCES


James C. Christensen is a research psychologist with the Air Force Research Laboratory. He received his PhD in cognitive psychology from The Ohio State University in 2008.

Justin R. Estep is a research biomedical engineer with the Air Force Research Laboratory. He received his BSE in biomedical engineering from Wright State University in 2006.

Date received: June 6, 2012
Date accepted: January 7, 2013