This article was downloaded by: [University of Central Florida]
On: 10 June 2014, At: 06:39
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK


Ergonomics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/terg20

# Task partitioning effects in semi-automated human-machine system performance

P.A. Hancock [a]
[a] Department of Psychology , University of Central Florida , Orlando , FL , 32826 , USA
Published online: 10 Sep 2013.

To cite this article: P.A. Hancock (2013) Task partitioning effects in semi-automated human-machine system performance, Ergonomics, 56:9, 1387-1399, DOI: 10.1080/00140139.2013.816374

To link to this article: http://dx.doi.org/10.1080/00140139.2013.816374

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

*Ergonomics*, 2013
Vol. 56, No. 9, 1387–1399, http://dx.doi.org/10.1080/00140139.2013.816374

Taylor & Francis
Taylor & Francis Group

# Task partitioning effects in semi-automated human–machine system performance

P.A. Hancock*

*Department of Psychology, University of Central Florida, Orlando, FL 32826, USA*

(*Received 10 September 2012; final version received 3 June 2013*)

Twelve professional pilots performed a flight simulation consisting of three component sub-tasks: (i) tracking, (ii) monitoring and (iii) targeting, respectively. The targeting sub-task required (i) target identification, (ii) weapon selection and then (iii) weapon release. Pilots performed in a fully manual condition, a partial automation condition or a fully automated condition. Automated assistance was provided for the targeting sub-task only, while tracking and monitoring sub-tasks were always performed manually. During full automation, the computer located the target, identified it and released the appropriate weapon without any pilot input. During partial automation, the computer located and identified the target while the pilot retained final control over weapon release. Significantly higher levels of tracking error distinguished manual from both automated conditions and also between the two levels of automation. Monitoring response times were also sensitive to the degree of automation engaged, with the partial-automation condition exhibiting faster responses than full automation. Findings support a design principle in which pilots retain control over final weapons release directly on the basis of objective performance outcome. These collective results support the contention that effective and principled task-partitioning should represent a central strategy for the evolution of complex human–machine systems.

**Practitioner Summary:** Advantages of partitioning tasks between human and automated control are contingent upon the overall context of performance and the actual way the partitioning is accomplished. Simple algorithms, for example, automate on every feasible occasion, are poor design heuristics and may even prove actively harmful to overall response capacity. Transitioning humans from active controllers to passive monitors can be a problematic design choice, especially when that individual is socially deemed to retain overall responsibility for ultimate system effects in the real world.

**Keywords:** degree of automation; pilot control; task partitioning

Downloaded by [University of Central Florida] at 06:39 10 June 2014

## Introduction

As we make progress towards fully adaptive human–machine systems, a central concern that still remains to be resolved is the appropriate degree of task partitioning to implement (Phatak and Bekey 1969a, 1969b; Sheridan 2002). Contemporary research has shown the advantage of maintaining at least some degree of operator awareness of the various sub-tasks, even if they are performed in a fully automated mode (Parasuraman and Riley 1997). For example, Parasuraman, Mouloua, and Molloy (1996) demonstrated that a task that is completely automated cannot simply be 'forgotten' but is in reality itself transformed into an additional, obligatory monitoring task. Unfortunately, humans have proved to be relatively deficient at such prolonged monitoring and any design rationale that advocates wholesale automation simply reduces operators to the monitors of multiple displays; a task which they are poorly equipped to perform (Bainbridge 1983; Hancock 1991, 2013). However, if automation does not meaningfully reduce operator load, or merely transforms it into other more diffuse forms, then simple indiscriminant automation can actually prove harmful to overall system performance. Thus, there is a need to take advantage of the opportunities offered by adaptive automation but not exploit this to the extent that overall system performance is degraded (Billings 1992; Rouse 1994; Sheridan 2002). Balancing these advantages against potential disadvantages provides a fundamental rationale for considering how the transitions into and out of automation should occur (Hancock 2007b) as well as what levels of task partitioning are viable between the operator and the machine. And it should be noted that some design conceptions advocate for a more holistic approach to human-in-the-loop system operation such that the human retains a more general picture of overall operations and no individual tasks are pruned off for completely automated completion (see e.g. Hoc, Young, and Blosseville 2009).

A number of strategies for dividing up sub-components of tasks, subsumed under the more general principles of function allocation (Birmingham and Taylor 1954; Chapanis 1965; for a review see Hancock and Scallen 1998), have previously been offered. For example, tasks can be shared in time. This multiplexing strategy recalls the earlier notion of complementarity (Jordan 1963) and has been investigated in work on automation 'cycling' (Hancock, Scallen, and Duley 1994; Scallen, Duley, and Hancock 1996). What makes such cycling a problematic long-term strategy are the 'hand-over' costs when manual control is assumed, relinquished or resumed. Given that the completely smooth transition of seamless

*Email: peter.hancock@ucf.edu

© 2013 Taylor & Francis

Downloaded by [University of Central Florida] at 06:39 10 June 2014

task switching could be achieved however, task cycling may prove an efficacious solution to the partitioning problem. In contrast to multiplexing, a second alternative is the spatial partitioning of tasks (e.g. Morrison, Cohen, and Gluckman 1993). Here, some salient features of the overall demand serve to divide the task itself into purportedly ‘natural’ and putatively discrete components. Then, some parts can be accomplished manually while others are handled by various degrees of automation. Some progress towards achieving this goal through the use of computational models has recently been advanced (Bolton, Siminiceanu, and Bass 2011; Kaber and Kim 2011).

One example of purely spatial task division could be flight control. While the automated system could deal with vertical control, represented by altitude above ground level, the pilot could still retain manual control over horizontal positioning. Such partitioning may well prove particularly useful in certain operations such as ‘nap-of-the-earth flight’. Automatic maintenance of the aircraft at a particular height above the ground could then free the pilot for other critical tasks such as weapon release, and such forms of control are, of course, now entering the realm of unmanned drone operations as well as modern ground vehicle control. In the air, however, the pilot still has to monitor the automated tracking system to make sure it is performing as required, especially as altitude control is an absolutely critical task (Hancock 2007a, 2007b). These questions of task parsing and division persist beyond traditional aircraft then to more contemporary issues involved in many unmanned systems (de Visser et al. 2008). The issue proliferates well beyond aviation since there are most relevant applications in numerous additional realms concerning, for example, ground vehicle control (Tsimhoni and Green 2001), control of assistive technologies (Simpson and Levine 1999) as well as more general perspectives that make clear such partitioning is a viable strategy across a wide spectrum of real-world conditions (Eischeid, Scerbo, and Freeman 1998; Nechyba and Xu 1997; Von Hippel 1990).

In the *cited aviation* example, the flight task is able to be divided ‘naturally’ into the two components of vertical and horizontal control. This division is based upon the physical nature of two dimensional tracking that also underlies the recent innovations in longitudinal control in advanced ground vehicles (Chen and Wang 2011). However, in the wider spectrum of complex systems, the appropriate spatial division of tasks is often not so readily apparent. Often ‘natural’ divisions are neither intuitive nor necessarily beneficial. Nowhere is this question of task division and performance responsibility brought more to the fore however, than in military aviation and in the vital decision to release or not release ordnance. Given a number of recent experiences in contemporary conflicts, the decision to fire or to withhold fire is now more than purely a tactical one. Indeed, as we have seen, it can rapidly burgeon into an implicit political decision (Bruxelles S.D. 2010 “Coroner Criticises US as He Gives ‘Friendly Fire’ Inquest Verdict” (*The Times*, April 28, 2010); Klein 2003; Snook 2000). This is so since modern improvements in weapons capacity now most frequently assures target destruction.

Founded upon both these theoretical and practical premises as well as the totality of the prior literature concerning function allocation and task partitioning (for a review see e.g. Hancock and Scallen 1998; and see also de Winter and Dodou 2011; Parasuraman, Mouloua, and Molloy 1996), the present experiment sought to address the question of task partitioning through the examination of part-task and full-task automation of this particularly critical weapons release sub-task during a simulated flight scenario. Specifically, the present procedure contrasted full automation of this specific sub-task in which the system detected a target, selected a weapon and fired that weapon, with a comparable, partial automation condition in which the system detected the target and selected the appropriate weapon but required the pilot to assure that a correct target had been identified and the appropriate weapon had been selected before firing. All of the time when the weapons sub-task was in progress pilots, explicitly remained in active control of the flight path of the aircraft and also remained responsible for monitoring aircraft systems status. Thus, the degree of automation here referred only to the one critical sub-task and so the partitioning referred only to a nominal one-third of the overall imposed demand. Such conditions have only been evaluated relatively rarely, even in the extensive literature on human–automation interaction.

The present experiment was one in a programmatic series which evaluated the effects of automation transitions and task-partitioning on response capacity when interacting with a complex operational system. The specific hypotheses for the present experiment were twofold. First, it was proposed that overall performance would be facilitated in trials in which any form of automation was employed; this when compared to the control trials in which all sub-tasks had to be accomplished manually. This hypothesis addressed a general task loading effect. The second hypothesis was that full automation would result in superior performance on the remaining sub-tasks, compared to both the part automation and the fully manual condition. However, a caveat to this hypothesis was that the pilots would experience greater cognitive workload in full, as compared to part-task automation as a result of their necessity to monitor the accuracy of automated weapons release. This proposition derives from previously suggested stress effects which increase in circumstances where an individual bears responsibility for an action but cannot affect immediate control over that action (see Hancock and Warm 1989; Karasek and Theorell 1990). Confirmation of this latter hypothesis would be indicative of elevated workload due to responsibility without authority in the full automation condition compared to the increased authority in the part-task automation, although the latter is actually accompanied by a slight increase in the necessity for active control. Finally, it was hypothesised that the demands of display scanning would modify these loading effects. This influence would be contingent upon each different display assemblage associated with the different sub-task combinations in the different levels of automation.

## Experimental method

### *Experimental participants*

To evaluate the above propositions, 12 experienced pilots (eleven males and one female) volunteered to participate in the present experiment. The pilots were recruited locally, partly from a group of professional participants undergoing advanced training at the Northwest Air Training Company as well as volunteers from the local professional flight community. Their experience ranged from General Aviation with flying time of just under 100 h in Visual Flight Rule conditions to Commercial Aviation with up to 5500 h flying time under Instrument Flight Rule restrictions. Only one of the pilots had any professional experience with fly-by-wire aircraft while the majority was familiar only with traditional forms of aircraft automation. The mean flight experience across the group was 1183 h. Although this range of experience, especially with the particular form of military aviation task to hand, might seem of concern, it has previously been shown that results from this type of sample range do generalise to other specific piloting tasks using differing forms of adaptive automation (and see Hancock and Smith 2007).

### *The STARFIRE test facility and sub-task description*

The present experiment was conducted in a single-seat, flight simulator. The body of the aircraft in which the pilot sat was a purpose-converted Agwagon which presented a large screen in the centre of the visual field and three, in-cockpit video display terminal (VDT) displays on which different information could be presented. In the present experiment, the STARFIRE test simulation was employed. This facility has previously been described in detail (Hancock and Scallen 1997). However, briefly, STARFIRE presents three component sub-tasks associated with flight, see Figure 1. The most evident of these flight-related sub-tasks is the tracking requirement which presents a *highway in the sky* that participants are required to follow. They were instructed to minimise their flight tracking error around a desired path, which is shown by the central ball in the ladder of the primary flight display. A second task, which is illustrated on the upper left side of Figure 1, represents flight systems monitoring. This requires the pilot to constantly assess the state of a series of five lights and four

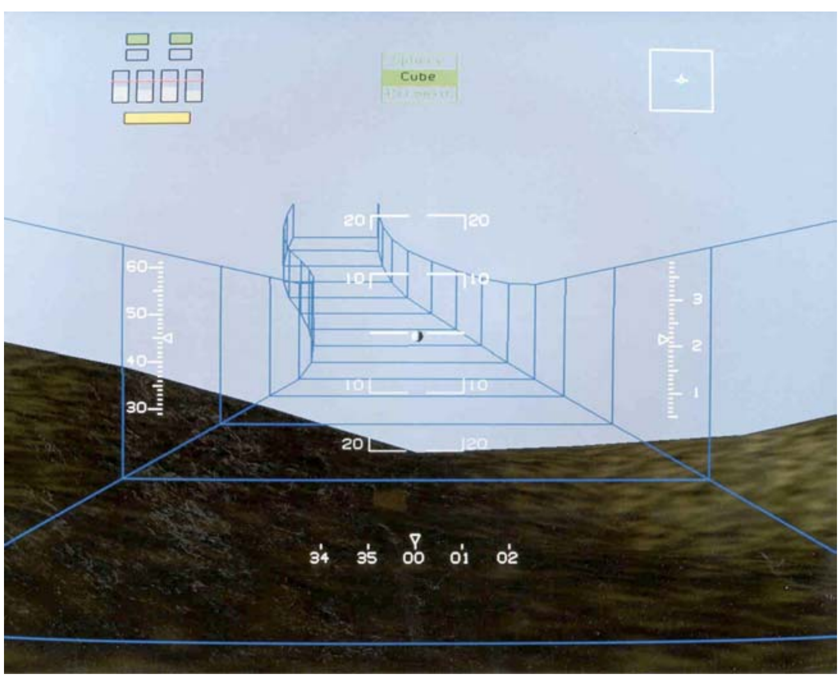


Figure 1. The STARFIRE display configuration during one of the automated trials. The monitoring lights and gauges are shown at upper left. Tracking automation is shown as being active as represented by the small box at upper right and performance is optimised as shown by the small ball at the centre of the pitch ladder. This form of tracking automation was not used in the present experiment. The targeting task, in the upper centre, is active and has selected a Cube target as shown. The target itself appears on the surface just above and to the left of the 00 direction indicator at the base of the HUD.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

Downloaded by [University of Central Florida] at 06:39 10 June 2014

gauges. The pilot is instructed to return these to their normal status whenever they deviate from an acceptable state. This is accomplished by activating the appropriate keys on an associated keypad. The third and final sub-task requires the pilot to search the terrain below them for pre-specified targets which appeared aperiodically. After recognising the presence of a target (i.e. the fact that a target is on the ground in front of them), the pilot then had to identify the specific configuration of the target (i.e. whether it was a cube, a sphere or a pyramid). They are then required to select the appropriate response (i.e. the weapon that would destroy that particular target). Each of these sub-tasks can be performed under full manual control or any of a number of different levels of automation. Further, all of the tasks could be partitioned so that computer and human could share the responsibility for their execution. In the present experiment, partitioning was engaged for the targeting task only.

### *Experimental conditions and design*

The current experimental design was a completely within-participants one. Pilots experienced conditions involving the full manual control of all of the sub-tasks on trials 1, 3 and 5. These respective trials were each of 4-min duration. Trials with either level of targeting task automation lasted for 6 min and occurred as trials 2 and 4 in the sequence. This regimen was imposed in order to balance the total time in manual control versus that in any form of automation, and reflected a further interest in the transition process both to and from automated control (Hancock 2007b). The order of the administration of different levels of automation was counter-balanced across participants. There were two forms of such automation, namely full and partial automation. In full automation, the computer located the target, identified that target and then released the appropriate weapon without any required input from the pilot. Selection of the duration of manual versus automated trials and the number of events in the various trials was based on balancing observations on response across the conditions for statistical comparison purposes. In the partial automation condition, the computer also located and then identified the target but the release of the weapon in this circumstance remained under direct pilot control. In order to uphold the integrity of the multi-task system, pilots were instructed that they still needed to monitor the targeting task even when it was in the full automation mode. Pilots were further instructed that upon completion of the full automation target task trial, they would be asked *How many surface targets did the automated system correctly destroy?* This question was employed specifically to prompt pilots to monitor full target automation so that this part of the overall effort could not simply be neglected.

Each manual trial contained eight system monitoring deviations and each automated trial contained 12 such monitoring deviations. System monitoring events were therefore scripted to occur randomly but on an average rate of two per minute in both manual and automated trials. A monitoring response was defined as the depression of the appropriate response switch after identification of the system monitoring display stimulus. Light and gauge deviations remained in abnormal status until the response switch was depressed or until the next deviation occurred. In the latter situation, the pilot was assigned a missed response.

In respect of the target task, every trial contained one target per minute but these did not necessarily appear at an even rate of one per each 60-s period. Thus, each manual trial contained four surface targets in total and each automated trial contained six surface targets. Of the 12 total targets there were four spheres, four cubes and four pyramids. It should be again emphasised that target task data were collected in manual trials and during the trial where the target task was partially automated. However, during the fully automated target task trial, the only data collected were the pilots' response to the question, *How many targets were correctly destroyed by the system?* An automated target task failure (where the system selected an incorrect weapon) was introduced for the last target in each of the two automated target task trials. Automation failures were introduced to ensure the pilot did monitor the automated task and thus maintained the overall integrity of the three part multi-task environment.

Since trials employing either form of automation were 6 min in length and contained 12 system monitoring deviations and 6 surface targets, and since manual trials were 4 min in length and contained eight monitoring deviations and four targets, the summed manual and automated trials were equated for each of the total time on task (12 min), the total number of monitoring deviations (24) and the total number of surface targets (12). The distribution of system monitoring deviations, and target types, within each trial was also equated for the three manual and two automated episodes.

### *Experimental procedure*

After first completing the informed participation form, pilots were asked to complete the Profile of Mood States (POMS) assessment instrument (McNair, Lorr, and Droppleman 1992) to assess their baseline mood state. This instrument has been previously employed in relation to assessing operator response to differing forms of part-task automation (Hancock 2007b). Pilots were then given instructions on the procedures and goals associated with each sub-task. The automated targeting sub-task and its partial form were then described and demonstrated. Pilots were allowed to practice on each and all of these component sub-tasks both singly and in combination to their own satisfaction. Prior to the flight performance, they also completed the Subjective Workload Assessment Technique (SWAT) card sort procedure (Reid and Nygren 1988) and then the experiment consisting of the five trial periods began. Before the start of the specific trial employing the fully automated targeting task, the pilots were reminded that they would be queried concerning the number of targets correctly identified by the system. At the end of each trial, the pilot was asked for their concluding SWAT response. Following the termination of all of the experimental trials, participants completed a post-test version of the POMS questionnaire. They were then de-briefed, thanked and were able to leave the facility.

### *Subjective responses*

Detailed information as to the composition and use of the POMS instrument is readily available in the literature (McNair, Lorr, and Droppleman 1992). In the present experiment, the modified 'right now' instructions were used for the post-test assessment. This modified instruction set was employed in order to make the test more sensitive to the specific, mood state changes associated with the experimental manipulations. The SWAT is also a common measurement instrument for subjective response and has been described in full detail elsewhere (see Reid and Nygren 1988). Here, SWAT responses were taken immediately after each of the five trials and these mean values were the subject of analysis. Since objective measures of performance and subjective perception of the task demand are not always coincident (Hancock 1996; Yeh and Wickens 1988), the present subjective measures sought to evaluate pilots' perceptions of the task partitioning effect alongside their objective performance response. Given previous positive evidence of change in subjective workload in relation to automation manipulation (Scallen and Hancock 2001), it was hypothesised that pilots would be sensitive to the change in the automation level with respect to the targeting task, and such effects would be evident in scores from both the POMS and SWAT scales.

### *Dependent performance measures*

The primary measure of performance on the tracking sub-task was root mean square (RMS) error calculated around the three-dimensional performance centerline (Miyake, Loslever, and Hancock 2001). System monitoring was measured in terms of the response time (RT), the overall percentage of responses made and the percentage of those responses which were correct. Targeting performance was measured by the percentage of responses made, the percentage of those responses which were correct and the mean time for response. Cases of missed targets and false alarms were also recorded. However, no direct targeting performance data were collected during the trial in which the task was fully automated. To reiterate, the first hypothesis anticipated a task loading effect, while the second hypothesis predicted an improvement in overall objective performance with complete automation, but increase in subjective load in this latter condition since weapons release was in this circumstance removed from the pilot's direct control.

## Experimental results

### *Tracking performance*

With respect to mean RMS error (RMSE), three separate analyses were performed. The first compared grouped manual (mean performance for all three manual trials) versus grouped automation (mean performance for the two target task automated trials) which looked to evaluate general variation in tracking due to any form of automation. The *t*-test indicated a significant difference ($t(11) = 6.876$, $p < 0.01$) between manual (i.e. 1.19) and automated (i.e. 1.37) conditions. This outcome confirmed the advantage that degrees of automation can bring, especially in complex, multi-task operations. Since pilots treat tracking as of primary importance, the significant improvement recorded here provides strong evidence that overall task load is reduced in the respective automation conditions. Task partitioning here therefore fulfills the first mandate of automation which is to improve performance on those tasks which are not subject to automation.

Having confirmed this foundational outcome, a secondary analysis of the tracking data was conducted on only those trials involving manual control. The mean RMSE value for each of the three manual trials was entered into a repeated measures ANOVA. This analysis elicited no significant effects ($p > 0.05$), showing that tracking did not change with time-on-task. This indicates that pilots, having practiced to asymptote before the experimental procedure retain stable levels of manual control across the duration of the experiment. Consequently, change in performance as a function of differing levels of automation is attributed to differences in task configuration and not outside influences such as learning or fatigue (see Matthews et al. 2012).

Having established the efficacy of targeting automation on tracking and also having shown that tracking was stable across sequential manual trials, a third analysis was conducted which compared tracking performance in the two trials

Downloaded by [University of Central Florida] at 06:39 10 June 2014

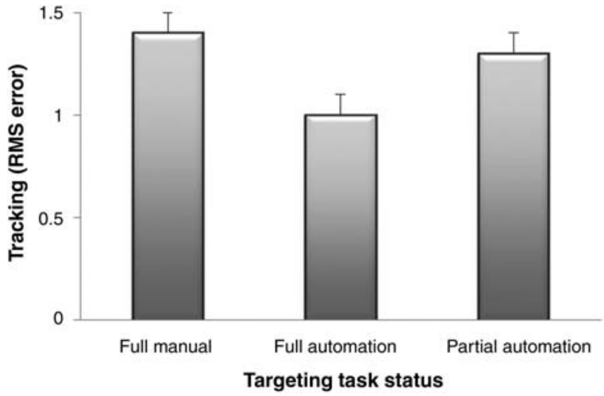


Figure 2. Tracking performance for the fully automated, partly automated and manual trials, with standard error bars shown. The figure displays tracking performance for the two trials in which tracking and system monitoring tasks were performed manually while the target task was either partially automated or fully automated. Data indicate that tracking performance was superior when the target task was fully automated. Furthermore, tracking performance in each of the two automated trials was superior to the performance across the summed manual trials.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

involving different degrees of automated targeting. Mean RMSE for trials employing either full or partial automation, respectively, was compared using an alpha-adjusted matched-pairs *t*-test. This showed a significant difference between the two conditions ($t(11) = 2.246$, $p < 0.05$). Here, fully automated targeting proved superior to partially automated targeting. Next, tracking summed across the manual trials was compared in a pairwise manner to each of the two automated trials. Paired *t*-tests indicated that tracking in manual trials was significantly different from both partial automation ($t(11) = 3.571$, $p < 0.01$) and full automation ($t(11) = 7.197$, $p < 0.01$) (see Figure 2). While the differences between these respective conditions may seem relatively small, there are many flight contexts, such as *nap-of-the-earth* operations or altitude keeping in busy airspaces, where even small errors in flight control accuracy can be disastrous.

### *System monitoring performance*

For systems monitoring, data were collected for each light and gauge deviation. The mean percentage response rate and mean RT were calculated for each of the five trials. In the present experiment, false alarms proved negligible as the sample of professional pilots responded to essentially all monitoring deviations and made virtually no false alarms. These respective ceiling and floor effects for the percent responses and false alarm data precluded meaningful signal detection analyses. Mean monitoring RT was calculated for each pilot in each trial and the initial analysis conducted concerned grouped RTs in the manual trials versus grouped RT's in the target task automated trials. The relevant *t*-test indicated no significant differences between these two groupings ($p > 0.05$). This appears to be because the pilots exhibited near perfect performance in the full multi-task condition already. RT's comparing each manual trial also showed no significant effects ($p > 0.05$). As with the comparable analysis for the tracking this outcome confirms that monitoring performance level stayed stable across the repeated manual trials. RT's for the two trials involving automated target task were then analysed using a matched pairs *t*-test. This analysis sub-served the main purpose of the present experiment, i.e. examining performance differences associated with differing levels of part-task automation associated with the overall targeting task. This separate *t*-test did indicate a significant difference ($t(11) = 2.275$, $p < 0.05$) in which the partially automated target task condition exhibited significantly lower (i.e. better) RTs (i.e. 3.28 s vs. 3.81 s, respectively, and see Kaber, Onal, and Endsley 2000). Analysis was then conducted on RTs summed for the manual trials compared to either form of automation; a comparison which elicited no significant differences ($p > 0.05$). Thus, retention of the weapons delivery function did not degrade speed of monitoring response. Further, the necessity to still identify targets during full automation had no differential influence compared to the full manual condition (and see Hancock 2007a, 2007b).

### *Target task performance*

The experimental design precluded sampling of objective target sub-task performance in the fully automated condition. Therefore, the present analysis involved comparing targeting sub-task performance across the manual trials with the trial in which partial automation occurred. An ANOVA was performed on the percentage of targeting responses across sequential manual trials and no significant effects were observed. This further confirmed stability of performance and the efficacy of

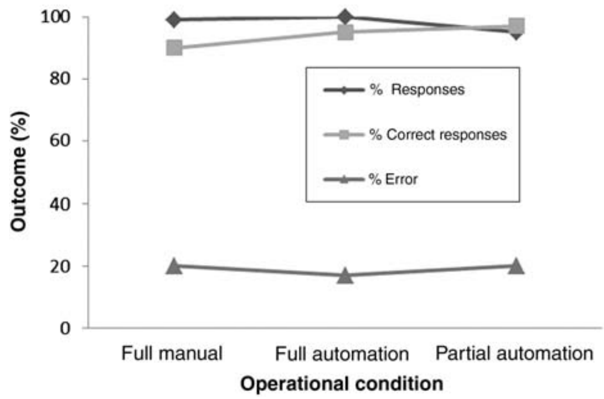


Figure 3. Levels of target task performance for the grouped manual trials, the trial employing partial automation of the target task and the trial employing full automation of the target task. Data indicate only slight differences between the conditions, which are assumed to be representative of an intrinsic ceiling effect. The data for the full automation condition are largely representative of system performance. The salient performance measure in this condition is the error metric that reflects pilots' response to the question 'How many targets were correctly destroyed by the system?' This, however, contains a form of memory demand that the immediate responses of the other two conditions do not. Despite this potential issue, the present pilots were very accurate in their recall of the number of targets destroyed in the full automation condition.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

the *ad libitum* regimen of practice. A subsequent analysis was performed on the percentage of correct responses which again elicited no significant differences. The absence of any significant difference in the targeting sub-task is thought to be due to the uniformly high level of pilot performance; essentially a ceiling effect. The relative infrequency of targeting events here (which does however reflect an appropriate ecological validity of the rarity of targeting in the real-world) means that opportunities to 'miss' the targets themselves were infrequent. The uniformly superior performance level attests to the importance which pilots (appropriately) attached to this critical task. In real-world circumstances, the appearance of a target is even rarer than was enacted in the present procedure. Thus, real-world targeting is much more akin to a full sustained attention task (and see Hancock 2013). However, only a limited number of targets had to be responded to and this identification problem is also representative of real-world demand. In the fully automated condition, pilots were uniformly observant, even correctly reporting on the interpolated targets where incorrect automation identification occured. This raises the issue of ecological validity and the generalisation of results, as discussed later in discussion.

To achieve a full description of response, targeting performance was represented by three specific metrics: (i) percentage of responses; (ii) percentage of correct responses and (iii) percentage of error. This latter value was calculated by dividing the number of targets involving an error by the total number of surface target presented; where an error was considered as any incorrect weapon fired *or* any missed target. Data in the full automation condition simply describe system behaviour. However, percentage of error here does reflect pilot errors in their response to the question concerning the number of targets correctly destroyed by the automated system. Pilots were asked this question at the end of the trial employing fully automated target task, so any failure to note incorrect system performance (and there was only one such occasion) meant that the fully automated condition did not always receive a 100% score on this final measure. These three different descriptive levels of performance are represented in Figure 3.

### *Subjective assessments*

Data for the SWAT measures of time load, stress level and mental effort were collected at the end of each of the five sequential trials. Response values, banded workload scores and specific workload scores were calculated for each measure, for each trial and subjected to analysis, which revealed no significant effects for any measure. Similarly, the POMS questionnaire scores were collected for the pre- and post-experimental sessions and these data also showed no significant differences. These results ran counter to the hypothesised pattern. They imply that there is no difference in subjective workload for the overall task demand since there was no main effect for automated versus manual conditions. They further show no difference between levels of automation also as hypothesised (and see Scallen and Hancock 2001). As the results ran counter to the trend of previous findings, they are of particular note and perhaps one possible explanation lies in the display augmentation to tracking here which may have diminished the level of momentary control demand in the STARFIRE ensemble as compared to, for example, earlier experiments using the MINSTAR facility (Hancock 2007a). The postulation that display design itself can mitigate performance variation and modulate experienced workload argues for continuing efforts to research the elusive notion of ‘optimal’ design (Hancock 2013; Hancock and Szalma 2003; Vicente and Rasmussen 1992). A second issue concerns the limited time over which any such workload variations are permitted to develop here. It may well be that the short exposures curtailed the opportunity to observe developing differences, since there is an intrinsic problem of memory-based recall in most subjective measures anyway. The null finding may reflect a coding in memory concern upon which there have been interesting recent developments (see Redelmeier and Kahneman 1996; Redelmeier, Katz, and Kahneman 2003).

Downloaded by [University of Central Florida] at 06:39 10 June 2014

## Discussion

The results of this experiment first confirm and then extend previous findings concerning part-task automation (cf. Endsley and Kaber 1999; Kaber and Endsley 2004; Scallen and Hancock 2001; Zhi-Gang, Macwan, and Wieringa 1998). While the changes in performance associated with partial versus full automation were not extensive, it should be recognised that in the present experiment there were only rather limited procedural differences in the responses required of the pilot by the targeting task. However, prior to discussing the task partitioning effects, it is important to emphasise the general advantage for the automated versus manual conditions. The present results show that the automation performance advantage is most evident in the remaining and centrally important tracking task which pilots, not unnaturally, treat as their highest priority. In both automation conditions, the pilot still retained the responsibility to either fire the correct weapon or to assure that the system had fired the correct weapon. These residual monitoring duties might have added to the overall load, but as shown by the tracking error findings and the absence of any change in level of subjective workload, it is clear that automation in its two respective forms proved beneficial to overall performance. This outcome therefore represents further support for the general case for the utility of automation as a general design strategy in human–machine systems (Farrell and Lewandowsky 2000; Hancock and Chignell 1988; Hancock, Chignell, and Loewenthal 1985). There are, however, limitations to interpretations placed on the present results. First, there were a relatively limited number of pilots, and their previous experience with advanced automation was individually unequal. In general, pilots with extensive experience of differing automation configurations are difficult to solicit for voluntary participation and the present pattern of responses must be considered with this level of pilot familiarity in mind. Further, the efforts to ensure ecologically reasonable conditions in terms of target frequency resulted here in ceiling effects which may have masked more subtle performance trends in longer exposures or with higher stimulus presentation frequencies. This issue, however, is a perennial one which faces all researchers who seek to make their experimental environments as close as possible to real-world conditions. Since unanticipated targets are, in the real-world, such a rare occurrence, the degrading influence of certain vigilance-related effects may not have been seen here because of the target frequency employed and the relatively short duration (e.g. Hancock 2013; although see Nuechterlein, Parasuraman, and Jiang 1983). Since the eventual purpose is to generalise to real-world operations, this trade-off is one which must be recognised and considered in the application of the present findings.

In the current experiment, the absolute level of RMSE was uniformly low but again this is not necessarily surprising for at least three reasons. First, the current display was purposefully created to be of particular assistance in supporting tracking accuracy. Indeed, as part of this general program of research, a conscious effort was made to improve the standard heads-up display (HUD) symbology with a *highway in the sky* augmentation, explicitly to achieve superior performance (see Figure 1). Second, pilots were given extensive tracking practice prior to engaging in the experimental conditions themselves. Third, the task of tracking is one with which pilots have much familiarity in relation to their prior professional experiences. Also, the frequency of automation errors was purposefully low to try to approximate real-world rather than artificial experimental circumstances. Given this, it is unsurprising that a very stable and low level of error was observed. This stability, especially across repeated manual trials, thus actually serves to emphasise the importance of the significant changes which did occur when different forms of automation were engaged. In the differential comparisons across the respective tracking results, the outcome showed that manual performance was poorer than that with partial automation which, in its turn, was poorer than that with full automation, a pattern echoed in some previous studies using the NASA Multi-Attribute Task battery (see Freeman et al. 1999; Prinzel et al. 2003). This pattern of performance initially appears to argue for full automation as the most effective design strategy. However, in any multi-task situation, using just one single sub-task performance metric as a basis for a general recommendation can prove highly misleading. In the present situation, the best monitoring task RTs occurred in the part-automation configuration. The question then becomes one of comparing how many milliseconds in monitoring task RT offsets what level of RMSE in tracking. This is not an easy trade-off to establish, especially since the value of this trade is always one that has to relate to actual, real-world performance constraints.

Due to the necessary design structure of the present experiment, a comparative, inferential analysis could not be performed on the targeting task results. However, the summed descriptive observations did suggest a trade-off between the two forms of automation. Here, percent correct responses were highest in partial automation condition while, in contrast, the sum of all errors, both incorrect response and missed signals was lowest in the full automation condition. There was also an apparent absence of a task demand effect on monitoring response. The present, experienced pilots had sufficient spare capacity so that there were essentially no missed signals or false alarms. The only significant differences that did occur were seen in momentary response latency. Here, there was an influence for the level of automation such that responses to partially automated situations were faster than those for full automation. This outcome does provide a degree of support for the contention of keeping the operator in the control loop where feasible (see also Parasuraman, Mouloua, and Molloy 1996). However, there is another important aspect of the present results. As improvements are made in the realm of automation and adaptive automation, the tendency is that the loading placed upon the human operator generally decreases. Indeed, the purpose of advanced interface design is often to simplify human performance requirements, sometimes to the extent that not only complacency but actual boredom sets in. When pilots are easily able to accomplish all tasks at a maximum performance level without disturbance to subjective state or diminishment from perfect performance, it may be that the task has been oversimplified. Thus an element involved in the search for a nominally 'optimal' design must include opportunities for engagement and exploration (Hancock, Pepe, and Murphy 2005), and not merely the progressive reduction in task demand.

The absence of effects in measures of subjective response is in contrast with previous findings (e.g. Hancock 2007a, 2007b). They are thus, initially at least, surprising. However, there may well be a straightforward reason for the absence of effect in the SWAT scores. This explanation lies in the respective length of the specific performance trials. Although the fully manual condition is evidently more demanding in terms of the required number of motor responses compared to either of the automation conditions, there was no difference in SWAT scores. However, the duration of each manual trial was only 4 min. The scores from these trials show no difference from those in the 6-min trials in either automation condition. The reason then for this null effect may well lie in the fashion in which subjective load is summed in memory (Hancock 1996). Since tracking demand drives both forms of trial, it may well be that the extra load in the manual condition is matched by the summation of load by the two extra minutes of each automation condition. In essence, given these duration differences, and given that task duration is known to systematically and linearly drive mental workload (Warm, Dember, and Hancock 1996), the null effect reported may perhaps mask an important difference. The idea that subjective workload is driven by crucial, discrete events, rather than being only a linear summation of perceived load over the interval of time has begun to be explored (Moroney, Biers, and Eggemeier 1995; Parasuraman and Hancock 2001; Svensson et al. 1997). This proposition accords with some recent work in the wider psychological sciences which also show that the phenomenological experience of pleasurable or painful events is calibrated by certain peak moments and extreme events rather than a summary record of everything that occurred in a particular interval of time (Redelmeier and Kahneman 1996). More importantly, there may be little difference in subjective apperception because the pilots were able to easily deal with the conditions which, nevertheless, prove much more demanding for relatively task-naïve individuals (see e.g. Hancock, Scallen, and Duley 1993).

We cannot, of course, divorce the present pattern of results from the physical configuration of the sub-tasks as they appeared on the presentation screen. As is evident from Figure 1, the central task, and one rated as the most important by pilots is the need for accurate and continuous tracking. The monitoring task appeared in the upper left and was visually peripheral to this central tracking location. Finally, the targets themselves appear on the terrain below the aircraft but since target detection is itself under computer control in both automation conditions, there is little intrinsic difference as to how the pilot scans the respective displays in these two circumstances. What is impressive is the maintenance of performance level for this task in the fully manual condition. Here, the scan pattern must certainly include frequent sampling of the terrain below. One interpretation of the pattern of findings reported can be derived in terms of the respective responsibility for individual task completion. In the manual condition, pilots must search for, identify and select and fire appropriate weapons as well as fly the aircraft and monitor its system status. This represents a challenging combination of demands but all of the sub-tasks are under pilot control. In partial automation, the pilot still has to fly the aircraft but receives help in the detection and identification phase of the targeting task. In such circumstances, the pilot does not have to spend so much time in the central locations, where both the tracking and the targeting task are located. Thus, the RT to peripheral monitoring deviations is enhanced by this freedom from the target search obligation. After all, since the system will detect and identify targets, the pilot simply has to confirm the veracity of this detection for weapons release and no weapons are fired without their consent. In the full automation condition, the pilot retains the responsibility to fly but the system may well detect, identify and fire weapons, even if the pilot's attention is engaged elsewhere. It is thus the dual responsibility to fly the aircraft and to monitor the autonomous actions of the system, which forces the pilots to retain their attention in the central location and therefore results in the relatively higher RTs to the peripheral monitoring events. These findings implicate an attentional strategy change and variation in display scanning as the level of automation is varied.

The conclusion here is that changing the form of automation results in differing forms of performance trade-off. There is a macro trade-off with respect to the other tasks in STARFIRE ensemble, i.e. tracking error is traded for monitoring RT.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

However, there is also a micro trade-off within the targeting task itself. The temptation is to draw upon a limited attentional construct to conclude that automated task partitioning allows for different attentional strategies under the respective task and interface constraints (Wickens and Carswell 1995). What does emerge is that pilots did monitor the completely automated display but engaged in a modified strategy compared to both manual and partially automated conditions. The conclusion of these findings is that even minor changes in the functionality of one component sub-task influences the performance profile across the whole suite of sub-tasks presented. It is anticipated that this finding will extrapolate across ever more complex interfaces, adapted automation tasks and real-world operational domains as explored below.

### Applications beyond aviation

It might seem initially as though the present results are very much constrained to aviation. However, this is far from the case. In reality, there is a tremendous range of potential applications for principles of human interaction with automated systems that have the capacity to change their sub-tasking profile. This is an important domain for the engineer and system designer since, in reality, there are now many and growing numbers of, and forms of, human-partially-in-the-loop automated and semi-automated systems. To the present at least, human beings remain intimately involved in most such systems, especially in their conception, design, fabrication and maintenance phases. At the very least, the product of most automated systems must eventually be communicated to a human operator in order to facilitate subsequent action. However, for any system that actually aspires to exclude all human interaction, there remain a vast range of other systems in which periodic interaction with human operators is the normal mode of operation. In actuality, many everyday technical systems have evolved towards this hybrid state (Hollnagel 2009). Recent specific exemplars include the remote control of unmanned aerial and ground vehicles (Taylor 2006), various manufacturing processes (Tolone et al. 1998), advanced commercial road vehicles (Shladover 1993) and general interaction with a variety of process control systems (Parasuraman, Sheridan, and Wickens 2000; Rouse 1981; Zachary et al. 1992). The general principles derived from the present experiment, e.g. the advisability to retain direct human control over the most critical task elements, are relevant to the conception, creation and operation of all of these forms of hybrid control systems. In respect to specific applications to industrial and manufacturing environments, there are direct implications for design evolution. For example, the issues of multi-tasking and task allocation and function reallocation are immediately applicable to all forms of human–robot interaction (Hancock et al. 2011; Keisel and Hinds 2004). Similarly, the parsing of operator tasks into active momentary control versus response to periodic system warnings is directly applicable to manufacturing process control (Hancock and Scallen 1996; Hwang and Salvendy 1988). Although the current specific pattern of results may be highly context-contingent to aviation displays, nevertheless certain general principles are likely to transcend the specifics of the present interface, for if they do not what are nomothetic design recommendations worth?

### Conclusion

It is important to make clear what is new in the present results. The use of highly motivated, experienced and well-practiced professionals showed fewer significant changes in response profile under discrete task partitioning than reported previously in studies using volunteer undergraduates (and for additional discussion see Hancock, Duley, and Scallen 1994). This perhaps warns that at least some, if not much, of our current body of understanding potentially derives from transient effects intrinsic to non-professional performers. The present regimen was chosen as a compromise between full ecological validity and the present need to evaluate performance over multiple events in experimental circumstances. Appropriate caution is thus needed when extrapolating from the present findings and it is especially important to understand here that the specific manipulations of automation status in the present work were not particularly pronounced. The essential difference between partial and full automation was in the final but critical decision to release the weapon. However, even with this relatively small degree of difference, significant effects on the overall performance profile were observed. This research provides direction to designers that the general conclusion that automation does influence performance positively is in itself a useful but insufficient conclusion. This is true because the form which that automation takes also affects how efficiency varies in a multi-task situation. Complex, multi-part tasks can thus be differentially designed to address such issues in current high-pressure, high-demand working environments. These results are encouraging for the continuing exploration of forms of task partitioning that differ from simple spatial or temporal divisions. The search for innovative ways in which candidate tasks might be partitioned promises to be a very fruitful avenue of future advance. Derived principles apply beyond any one specific domain from which they have been extracted. Indeed, these findings apply to all forms of hybrid human–computer control, whether such tasks are embedded in complex vehicle control, human–robot interaction, advanced manufacturing facilities and/or more general industrial process control.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

## References


Bainbridge, L. 1983. "Ironies of Automation." *Automatica* 19: 775–779.

Billings, C. E. 1992. *Aviation Automation: The Search for a Human-Centered Approach*. Mahwah, NJ: Erlbaum.

Birmingham, H. P., and F. V. Taylor. 1954. "A Design Philosophy for Man–Machine Control Systems." In *Proceedings of the IRE*, 1748–1758. New York: Institute of Radio Engineers.

Bolton, M. L., R. I. Siminiceanu, and E. J. Bass. 2011. "A systematic Approach to Model Checking Human Automation Interaction Using Task Analytic Models." *IEEE Transaction on Systems, Man, and Cybernetics: Systems and Humans* 41 (5): 961–976.

Chapanis, A. 1965. "On the Allocation of Functions Between Men and Machines." *Occupational Psychology* 39: 1–11.

Chen, Y., and J. Wang. 2011. "Adaptive Vehicle Speed Control with Input Injections for Longitudinal Motion Independent Road Frictional Condition Estimation." *IEEE Transaction on Vehicular Technology* 60 (3): 839–848.

de Visser, E. J., M. S. Cohen, M. Le Goullon, O. Sert, A. Freed, E. Freedy, G. Weltman, and R. Parasuraman. 2008. "A Design Methodology for Controlling, Monitoring, and Allocating Unmanned Vehicles." In *Third International Conference on Human Centered Processes (HCP-2008)*. Netherlands: Delft.

de Winter, J. C. F., and D. Dodou. 2011. "Why the Fitts List has Persisted Throughout the History of Function Allocation." *Cognition, Technology and Work*. doi: 10.1007/s10111-011-0188-1.

Eischeid, T. M., M. W. Scerbo, and F. G. Freeman. 1998. "The Effects of Task Partitioning and Computer Skill on Engagement and Performance with an Adaptive, Biocybernetic System." *Proceedings of the Human Factors and Ergonomics Society* 42: 133–137.

Endsley, M. R., and D. B. Kaber. 1999. "Level of Automation Effects on Performance, Situation Awareness and Workload in a Dynamic Task." *Ergonomics* 42 (3): 462–492.

Farrell, S., and S. Lewandowsky. 2000. "A Connectionist Model of Complacency and Adaptive Recovery Under Automation." *Journal of Experimental Psychology: Learning, Memory and Cognition* 26 (2): 395–410.

Freeman, F. G., P. J. Mikulka, L. J. Prinzel, and M. W. Scerbo. 1999. "Evaluation of an Adaptive Automation System Using three EEG Indices with a Visual Tracking System." *Biological Psychology* 50: 61–76.

Hancock, P. A. 1991. "On Operator Strategic Behavior." *Proceedings of the International Symposium on Aviation Psychology* 6: 999–1007.

Hancock, P. A. 1996. "Effects of Control Order, Augmented Feedback, Input Device and Practice on Tracking Performance and Perceived Workload." *Ergonomics* 39 (9): 1146–1162.

Hancock, P. A. 2007a. "The Effects of Automation Invocation Procedure and Dynamic Display Relocation on Performance in a Multi-Task Environment." *Transactions of the IEEE Systems, Man, and Cybernetics, Part A. Systems and Humans* 37 (1): 47–57.

Hancock, P. A. 2007b. "On the Process of Automation Transition in Multi-Task Human–Machine Systems." *IEEE Transactions on Systems, Man, and Cybernetics, Part A. Systems and Humans* 37 (4): 586–598.

Hancock, P. A. 2013. "In Search of Vigilance: The Problem of Iatrogenically Created Psychological Phenomena." *American Psychologist* 68 (2): 92–109.

Hancock, P. A., D. R. Billings, K. Olsen, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman. 2011. "A Meta-Analysis of Factors Impacting Trust in Human–Robot Interaction." *Human Factors* 53 (5): 517–527.

Hancock, P. A., and M. H. Chignell. 1988. "Mental Workload Dynamics in Adaptive Interface Design." *Transactions of the IEEE Systems, Man, and Cybernetics*, 18 (4): 647–658.

Hancock, P. A., M. H. Chignell, and A. Loewenthal. 1985. "An Adaptive Human–Machine System." In *Proceedings of the IEEE International Conference on Systems, Man,* and *Cybernetics*, 627–630. IEEE.

Hancock, P. A., J. A. Duley, and S. F. Scallen. 1994. *The Control of Adaptive Function Allocation*. Report HFRL, Nav-6, Naval Air Warfare Center, Warminster, PA, USA.

Hancock, P. A., A. Pepe, and L. L. Murphy. 2005. "Hedonomics: The Power of Positive and Pleasurable Ergonomics." *Ergonomics in Design* 13 (1): 8–14.

Hancock, P. A., and S. F. Scallen. 1996. "The Future of Function Allocation." *Ergonomics in Design* 4 (4): 24–29.

Hancock, P. A., and S. F. Scallen. 1997. "The Performance and Workload Effects of Task Re-Location During Automation." *Displays* 17: 61–68.

Hancock, P. A., and S. F. Scallen. 1998. "Allocating Functions in Human–Machine Systems." In *Viewing Psychology as a Whole: The Integrative Science of William N. Dember*, edited by R. R. Hoffman, M. F. Sherrick, and J. S. Warm, 509–539. Washington, DC: American Psychological Association.

Hancock, P. A., S. F. Scallen, and J. A. Duley. 1993. *Initiation and Cessation of Automation: Location Versus Configuration Change*. Report HFRL, Nav-3, Naval Air Warfare Center, Warminster, PA, USA, August.

Hancock, P. A., S. F. Scallen, and J. A. Duley. 1994. *Pilot Performance and Preference for Cycles of Automation in Adaptive Function Allocation*. Technical Report, Naval Air Warfare Center Warminster, PA, USA.

Hancock, P. A., and K. Smith. 2007. "A Distributed Air Traffic Information Display Simulator: Design and Results." *International Journal of Applied Aviation Studies* 7 (2): 232–243.

Hancock, P. A., and J. L. Szalma. 2003. "Operator Stress and Display Design." *Ergonomics in Design* 11 (2): 13–18.

Hancock, P. A., and J. S. Warm. 1989. "A Dynamic Model of Stress and Sustained Attention." *Human Factors* 31: 519–537.

Hoc, J.-M., M. S. Young, and J.-M. Blosseville. 2009. "Cooperation Between Drivers and Automation: Implications for Safety." *Theoretical Issues in Ergonomics Science* 10 (2): 135–160.

Hollnagel, E. 2009. *The ETTO Principle: Efficiency Thoroughness Trade-off*. Chichester: Ashgate.

Hwang, S-L., and G. Salvendy. 1988. "Operator Performance and Subjective Response in Control of Flexible Manufacturing Systems." *Work & Stress* 2 (1): 27–39.

Jordan, N. 1963. "Allocation of Functions Between Man and Machines in Automated Systems." *Journal of Applied Psychology* 47: 161–165.


Downloaded by [University of Central Florida] at 06:39 10 June 2014

Kaber, D. B., and M. R. Endsley. 2004. "The Effects of Level of Automation and Adaptive Automation on Human Performance, Situation Awareness and Workload in a Dynamic Control Task." *Theoretical Issues in Ergonomic Science* 5 (2): 113–153.

Kaber, D. B., and S-H. Kim. 2011. "Understanding Cognitive Strategy with Adaptive Automation in Dual-Task Performance Using Computational Cognitive Models." *Journal of Cognitive Engineering and Decision Making* 5 (3): 309–331.

Kaber, D. B., E. Onal, and M. R. Endsley. 2000. "Design of Automation for Telerobots and the Effect on Performance, Operator Situation Awareness, and Subjective Workload." *Human Factors and Ergonomics in Manufacturing* 10 (4): 409–430.

Karasek, R., and T. Theorell. 1990. *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life*. New York, NY: Basic Books.

Keisel, S., and P. Hinds. 2004. "Human–Robot Interaction." *Human–Computer Interaction* 19 (1–2): 1–8.

Klein, J. J. 2003. "Problematic Nexus: Where Unmanned Combat Air Vehicles and the Law of Armed Conflict Meet." Accessed September 2012. http://www.airpower.maxwell.af.mil/airchronicles/cc/klein.html

Matthews, G., P. A. Desmond, C. Neubauer, and P. A. Hancock, eds. 2012. *The Handbook of Operator Fatigue*. Farnham, Surrey: Ashgate.

McNair, D. M., M. Lorr, and L. F. Droppleman. 1992. *Manual for the Profile of Mood States*. Rev. Ed. San Diego, CA: EdITS.

Moroney, W. F., D. W. Biers, and F. T. Eggemeier. 1995. "Some Measurement and Methodological Consideration in the Application of Subjective Workload Measurement Techniques." *International Journal of Aviation Psychology* 5 (1): 87–106.

Morrison, J. G., D. Cohen, and J. P. Gluckman. 1993. *Prospective Principles and Guidelines for the Design of Adaptively Automated Crew-Stations. The Adaptive Function Allocation for Intelligent Cockpits (AFAIC) Program: Interim Research and Guidelines for the Appellation of Adaptive Automation*. Technical Report, NAWCADWAR-93931-60, Naval Air Warfare Center, Warminster, PA, USA.

Miyake, S., P. Loslever, and P. A. Hancock. 2001. "Individual Differences in Tracking." *Ergonomics* 44 (12): 1056–1068.

Nechyba, M. C., and Y. Xu. 1997. "Human Control Strategy: Abstraction, Verification, and Replication." *IEEE Control Systems* 17 (5): 48–61.

Nuechterlein, K. H., R. Parasuraman, and Q. Jiang. 1983. "Visual Sustained Attention: Image Degradation Produces Rapid Sensitivity Decrement over Time." *Science* 220 (4594): 327–329.

Parasuraman, R., and P. A. Hancock. 2001. "Adaptive Control of Mental Workload." In *Stress, Workload, and Fatigue*, edited by P. A. Hancock and P. A. Desmond, 305–320. Mahwah, NJ: Lawrence Erlbaum.

Parasuraman, R., M. Mouloua, and R. Molloy. 1996. "Effects of Adaptive Task Allocation on Monitoring of Automated Systems." *Human Factors* 38: 665–679.

Parasuraman, R., and V. Riley. 1997. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39: 230–253.

Parasuraman, R., T. B. Sheridan, and C. D. Wickens. 2000. "A Model for Types and Levels of Human Interaction with Automation." *Transactions of the IEEE Systems, Man, and Cybernetics* 30 (3): 286–297.

Phatak, A. V., and G. A. Bekey. 1969a. "Model of the Adaptive Behavior of the Human Operator in Response to a Sudden Change in the Control Situation." *IEEE Transactions on Man–Machine Systems* 10 (3): 72–80.

Phatak, A. V., and G. A. Bekey. 1969b. "Decision Processes in the Adaptive Behavior of Human Controllers." *IEEE Transactions on Systems Science and Cybernetics* 5 (4): 339–351.

Prinzel, L. J., F. G. Freeman, M. Scerbo, P. J. Mikulka, and A. T. Pope. 2003. "Effects of a Psychophysiological System for Adaptive Automation on Performance, Workload, and Event-Related Potential P300 Component." *Human Factors* 45 (4): 601–613.

Redelmeier, D. A., and D. Kahneman. 1996. "Patients' Memories of Painful Medical Treatments: Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures." *Pain* 66 (1): 3–8.

Redelmeier, D. A., J. Katz, and D. Kahneman. 2003. "Memories of Colonoscopy: A Randomized Trial." *Pain* 104 (1–2): 187–194.

Reid, G. B., and T. Nygren. 1988. "The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload." In *Human Mental Workload*, edited by P. A. Hancock, and N. Meshkati, 185–218. New York: North Holland.

Rouse, W. B. 1981. "Human–Computer Interaction in the Control of Dynamic Systems." *Computing Surveys* 13 (1): 71–99.

Rouse, W. B. 1994. "Twenty Years of Adaptive Aiding: Origins of the Concept and Lessons Learned." In *Human Performance in Automated Systems: Current Research and Trends*, edited by M. Mouloua, and R. Parasuraman, 249–255. Hillsdale, NJ: Erlbaum.

Scallen, S. F., J. A. Duley, and P. A. Hancock. 1996. "Pilot Performance and Preference for Short Cycles of Automation in Adaptive Function Allocation." *Applied Ergonomics* 26: 397–403.

Scallen, S. F., and P. A. Hancock. 2001. "Implementing Adaptive Function Allocation." *International Journal of Aviation Psychology* 11 (2): 197–221.

Sheridan, T. B. 2002. *Humans and Automation: System Design and Research Issues*. Santa Monica, CA: Wiley.

Shladover, S. E. 1993. "Research and Development Needs for Advanced Vehicle Control Systems." *IEEE Micro* 13 (1): 11–19.

Simpson, R., and S. P. Levine. 1999. "Automatic Adaptation in the NavChair Assistive Wheelchair Navigation System." *IEEE Transactions on Rehabilitation Engineering* 7 (4): 452–463.

Snook, S. A. 2000. *Friendly Fire: The Accidental Shootdown of US Black Hawks over Northern Iraq*. Princeton: Princeton University Press.

Svensson, E., M. Angelborg-Thanderz, L. Sjoberg, and S. Olsson. 1997. "Information Complexity and Performance in Combat Aircraft." *Ergonomics* 40 (3): 362–380.

Taylor, R. M. 2006. "Human Automation Integration for Supervisory Control of UAVs." In *Virtual Media for Military Applications, Meeting Proceedings RTO-MP-HFM-136, Paper 12*. 12, 1–10. France: Neuilly-sur-Seine.

Tolone, W., B. Chu, J. Long, R. Willhelm, Y. Peng, and A. Boughannam. 1998. "Supporting Human Interactions Within Integrated Manufacturing Systems." *International Journal of Advanced Manufacturing* 1 (2): 221–234.

Tsimhoni, O., and P. Green. 2001. "Visual Demand of Driving and the Execution of Display-Intensive in Vehicle Tasks." *Proceedings of the Human Factors and Ergonomics Society* 45: 1586–1590.

Downloaded by [University of Central Florida] at 06:39 10 June 2014

Vicente, K. J., and J. Rasmussen. 1992. "Ecological Interface Design: Theoretical Foundations." *IEEE Transactions on Systems, Man and Cybernetics* 22: 589–606.

Von Hippel, E. 1990. "Task Partitioning: An Innovation Process Variable." *Research Policy* 19: 407–418.

Warm, J. S., W. N. Dember, and P. A. Hancock. 1996. "Vigilance and Workload in Automated Systems." In *Automation and Human Performance: Theory and Applications*, edited by R. Parasuraman and M. Mouloua, 183–200. Mahwah, NJ: Lawrence Erlbaum.

Wickens, C. D., and C. M. Carswell. 1995. "The Proximity Compatibility Principle: Its Psychological Foundation and Relevance to Display Design." *Human Factors* 37 (3): 473–494.

Yeh, Y. Y., and C. D. Wickens. 1988. "Dissociation of Performance and Subjective Measures of Workload." *Human Factors* 30: 111–120.

Zachary, W., J. Ryder, L. Ross, and M. Z. Weiland. 1992. "Intelligent Human–Computer Interaction in Real-Time, Multi-Tasking Process Control and Monitoring Systems." In *Design for Manufacturability: A Systems Approach to Concurrent Engineering*, edited by M. Heldander, and M. Nagamachi, 377–401. Bristol, PA: Taylor & Francis.

Zhi-Gang, W., A. P. Macwan, and P. A. Wieringa. 1998. "A Quantiative Measure for Degree of Automation and its Relation to System Performance and Mental Load." *Human Factors* 40 (2): 277–295.

Downloaded by [University of Central Florida] at 06:39 10 June 2014