THEORETICAL ISSUES IN ERGONOMICS SCIENCE (IN PRESS)

A Proposed Psychological Model of Driving Automation

Neville A. Stanton and Mark S. Young
Department of Design
Brunel University
Runnymede Campus
Egham
Surrey
TW20 0JZ

This paper considers psychological variables pertinent to driver automation. It is anticipated that driving with automated systems is likely to have a major impact on the drivers and a multiplicity of factors needs to be taken into account. A systems analysis of the driver, vehicle and automation served as the basis for eliciting psychological factors. The main variables to be considered were: feedback, locus of control, mental workload, driver stress, situational awareness and mental representations. It is expected that by anticipating the effects on the driver brought about by vehicle automation could lead to improved design strategies. Based on research evidence in the literature, the psychological factors were assembled into a model for further investigation.

KEYWORDS: Automation, driving, feedback, locus of control, mental workload, driver stress, situational awareness and mental representations

1. Introduction

Whether we like it or not, automation is gradually taking over the driver's role. Full vehicle automation is predicted to be on British roads by 2030 (Walker, Stanton and Young, 2000). Whilst it is accepted that some drivers will still want to control their vehicles manually, many may be pleased to relinquish the role to automatic systems. Many of the computing technologies have been grounded in aviation systems (Billings, 1997; Stanton and Marsden, 1996), and technologies like adaptive cruise control are taking over from the driver already (Richardson, Barber, King, Hoare and Cooper, 1997). Adaptive Cruise Control (ACC) heralds a new generation of vehicle (Stanton, Young & McCaulder, 1997). ACC controls both speed and headway of the vehicle, braking with limited authority in the presence of a slower lead vehicle, and returning to the set speed when the lead vehicle disappears (Richardson, Barber, King, Hoare and Cooper, 1997). In this way ACC differs from traditional Cruise Control (CC) systems. In traditional cruise control, the system relieves the driver of foot control of the accelerator only (i.e., relieving the driver of some physical workload), whereas ACC relieves the driver of some of the decision making elements of the task, such as deciding to brake or change lanes (i.e., relieving the driver of some mental workload), as well as physical demands of accelerator control. Potentially, then, ACC is a welcome additional vehicle system that will add comfort and convenience to the driver (Richardson et al., 1997). However, certain psychological issues do arise when considering any form of automation and these need to be properly addressed to improve overall system performance. It is envisaged that although the ACC system will behave in exactly the manner prescribed by the designers and programmers, this may lead to some scenarios in which the driver's perception of the situation is at odds with the system operation (Stanton and Young, 1998). Indeed, even those developing the systems recognise that "*headway control raises the issue of whether the system matches the driver expectations with regard to braking and headway control*" (Richardson et al., 1997; p. 91).

Most of the literature on driver behaviour tends to be restricted to the examination of a very limited set of variables. Whilst it might be argued by the researchers that this focus is necessary in order to determine the importance of the interaction between two or three important variables, it does not constitute the complexity of interplay between variables in the world at large. Sophisticated experimental design and analysis methodologies are able to cope with this complexity however, to provide an understanding of multiple interacting variables necessary to develop a psychological model of driving with automation. The purpose of this paper is, then, to identify the relevant the psychological variables and, on the basis of the literature, propose a psychological model.

An in-depth analysis of the psychological factors associated with the operation of automated systems is required to enable recommendations to be developed. The pertinent factors were elicited from a systems model of the driver-automation-vehicle as shown in figure 1, discussed by Stanton and Marsden (1996). From this figure, some potential psychological constructs emerge.

FIGURE ONE ABOUT HERE

Figure 1. Information flow between driver, automatics and vehicle sub-systems (from Stanton and Marsden, 1996).

Most obvious is the issue of *feedback*, as can be seen as the information flows around the subsystems. Of particular interest is the role of feedback to the driver from the automated system. Typically, this tends to be poor (e.g., the lack of feedback from automated systems in the domain of aviation has been implicated in some incidents, Norman, 1990), because the automated systems do not require the feedback to function. Relatedly, the development of the driver's *trust* in the automated system may depend upon appropriate feedback. According to Muir and Moray (1996), the amount of feedback sought from an automated system by a human operator is directly related to the degree of trust they have in it to perform without failure. Passing control of the vehicle to a computer raises the issue of *locus-of-control* in the driver: does the driver feel that they, or the computer, is ultimately in control of the automobile? The degree to which a symbiotic relationship exists between the driver and the automatic system could determine how successful vehicle automation is perceived. One of the claims of all forms of automation is that the demands placed upon human operators will be reduced (Bainbridge, 1983). Therefore, the effects upon *mental workload* need to be considered. The workload literature suggests that there is an optimum level that leads to enhanced performance, demands greatly above or below this level could have a negative effect the driver (Young & Stanton, 1997), such as increased levels of *stress*. Matthews and Desmond (1996) have found that driver stress is an important factor in the drivers like, or dislike, of driving and is linked to the their experience of mental workload. One of the central concepts in driver automation seems to be the extent to which the driver is aware of the state of the automatic system, and the impact that has on the vehicle trajectory through the world. This concept is called *situational awareness* and has been the subject of research in the field of aviation (Endsley, 1995). Situational awareness depends, to a great extent, upon the development of an accurate model of the world, that enables information to be interpreted and predictions of future states made. Therefore, the role of *mental models* will also be considered. All of these factors are well established in the psychological literature but have yet to be fully explored with respect to vehicle automation.

2. Feedback

The significance of feedback to performance was recognised more than 40 years ago, when Annett and Kay (1957) wrote about 'knowledge of results' (KR). KR takes many forms, from a summary end score at the end of a task to augmented feedback during a task. This is analogous to the distinction between action feedback (immediate notification of the un/successful completion of an action) and learning feedback (more in-depth knowledge of one's performance during a trial, typically given through tuition). The former type of feedback predicts fast learning, but also fast decay of the relevant skill when feedback is removed; whilst the latter predicts slower learning and slower decay. The content of feedback can be about system actions and/or responses. Or it can be simply registering a user's input. Either way, KR has been popularly assumed to be a motivator by providing knowledge of goal achievement.

Welford (1968) reviews the early experiments into the effects of KR on performance. It was widely found that feedback is instrumental in skill acquisition. Performance that shows no change under control conditions can rapidly improve upon the introduction of KR. The same skill can then rapidly deteriorate when KR is removed, as people have a tendency to overcompensate for their own behaviour. However, this is not the whole story - most training regimes rely on feedback for instruction. These would ultimately fail when the apparent skills were applied in the field, when feedback is absent. Training can be effective if it uses less direct feedback, information that directs the participant to remember the feel of their actions, rather than simply the outcome. By utilising cues that are inherent in the task, the participant can then transfer these skills to the operational environment. If s/he comes to depend on extra cues that are available in training but not in the actual job, skill transfer will not occur.

There are also a number of experiments reviewed by Welford (1968) on the timing and quality of KR. Feedback only seems to be effective if there is minimal delay between action and feedback, and between feedback and the next trial. Any intervening activity will be particularly destructive to learning from feedback. Learning will also be disrupted if the quality of the information given is distorted or inaccurate. The key seems to be providing full and accurate information about the discrepancy between required and achieved performance. However, there is a danger of giving too much feedback, as superfluous information can be distracting.

More recently, however, Kluger and Adler (1993) found that feedback about the outcome of an action actually reduced motivation, whether the feedback came from another person or from a computer. It was also found in this study that feedback from another person can be detrimental to performance, although given a free choice people are more likely to seek feedback from a computer than another person.

It is well understood that KR is a crucial factor in the early stages of skill acquisition (Groeger, 1997). This has been applied to many diverse fields, from consumer products (Bonner, 1998) to aviation (White, Selcon, Evans, Parker and

Newman, 1997). In the latter study, it was found that providing redundant information from an additional source can actually elicit a performance advantage.

One study that is relevant to the driving domain examined the effects of feedback on performance of controlled and automatic tasks (Tucker, Macdonald, Sytnik, Owens and Folkard, 1997). Driving, as a highly skilled activity, is an archetypal automatic task. That is, it requires very little cognitive processing effort, and is very rapidly executed. In the Tucker et al. (1997) study, it was found that feedback can reduce error rates on tasks requiring controlled processing, however automatic tasks are resistant to the effects of feedback. Furthermore, a vigilance decrement was observed only in the controlled task, suggesting automatic responses do not suffer from such a decrement. This vigilance decrement was also unaffected by feedback.

Moving on to consider studies more specific to driving and automation, Duncan, Williams and Brown (1991) examined the performance of a group of normal (experienced) drivers with that of novices and experts (observers from the Institute of Advanced Motorists) on a subset of driving skills. They found that on half of the measured skills, the normals actually performed worst, with novices performing at a similar level to the experts. It was concluded that these skills were those which did not benefit from immediate, task-intrinsic feedback. For instance, a bad gear change is immediately recognised from the noise and movement of the car. However, failing to check a mirror at a junction may not necessarily result in any negative consequences. In short, the normal drivers had succumbed to a range of bad habits in the absence of learning feedback.

Fairclough, May and Carter (1997) examined the effects of continuous time headway (THW) feedback on following behaviour. Their results demonstrated that feedback reduces the percentage of time spent at low headways (i.e., <1s), particularly for those people who were habitual close followers.

3. Trust

The concept of trust has only been applied to technological devices relatively recently. Muir (1994) hypothesised that supervisory control is based on trust, such that intervention behaviour depends on the supervisor's level of trust in the automated system. By building on previous theories of interpersonal trust, Muir developed a model of trust in automated systems.

Muir (1994) reviews the literature on trust, and finds "surprisingly little research" on trust in other people. There are, however, a number of attempts to define trust, suggesting it is a multidimensional construct consisting of confidence in, expectations of, and reliance upon others. Muir (1994) finds that trust is commonly oriented toward the future, and always has a specific referent. She then settles on and cites the definition of Barber (1983), which refers to three specific expectations which underline trust: the persistence of the natural order, technically competent role performance, and the fiduciary obligations and responsibilities of others. A model of the development of trust in another person is then reviewed by Muir (1994), which is later adapted to trust in machines. The model is hierarchical, proposing that trust develops in three stages, such that the outcome at each stage determines later development. Predictability governs the first stage, the second is dependent on dependability, and faith dominates mature levels of trust.

A key element of the model when applied to automated systems appears to be predictability, such that supervisors more readily come to trust consistent systems. However, it is important to distinguish perceived trust from the objective trustworthiness or reliability of the system, as an imbalance can lead to mistrust or distrust. This is critical to the performance of the system, as distrust can lead to under-utilisation of the system (and hence suboptimal performance), whereas mistrust can lead to overreliance and the possibility of errors.

This model was later tested empirically (Muir and Moray, 1996). Subjective ratings of trust were gathered for a simulated milk pasteurisation plant. One of the best predictors of trust was found to be expectation of competence, although development of trust over time did not match the model predictions. Development of trust apparently follows a logarithmic function, with initial development being conservative, however this moves to extremes with further experience. It was also found that distrust can spread within components of a subsystem but not across different systems. The most significant practical finding, however, was that trust *determines* the use of automation - a trusted system will be used appropriately, however a system which does not foster trust will be neglected.

Lee and Moray (1994) extended this argument by considering the role of confidence in one's own ability. A process control simulation and subjective ratings of trust and self-confidence were used to determine that, in general, operators will use automation when trust exceeds self-confidence, however will revert to manual control when the opposite is true. Again, these effects were restricted to individual subsystems, and did not generalise across systems.

These ideas have been addressed by a few other authors. Parasuraman and his colleagues (Parasuraman, 1997; Parasuraman, Singh, Molloy and Parasuraman, 1992) considered the effects of automation on the human operator. Parasuraman et al. (1992) were interested in the concept of automation-related complacency, and developed a scale for attitudes toward automation which reflect a potential for complacency. Complacency was found to be comprised of trust, confidence, reliance, and safety-related complacency. It was not clear, however, whether complacency is situationally induced or an individual difference variable. It has been argued elsewhere (Kaber and Endsley, 1997) that inappropriate allocation of function can induce complacency, and complacency can lead to overtrust.

Parasuraman (1997) concentrated on the factors that can influence use of automation. Reliability of the system affects trust, which in turn interacts with self-confidence to determine reliance on the automation. Conversely, a propensity of false alarms can lead to distrust and disuse of an automated system. Overall, then, it seems that trust is an important factor in determining the extent to which an operator will exploit the automation at his/her disposal. There is some evidence to suggest that operators will not abandon using a system even when it is completely unreliable

(McFadden, Giesbrecht and Gula, 1998), however the majority of the literature indicates that trust, complacency, and task allocation should all be considered as important factors in the design of an automated system.

4. Locus of Control

Locus of control refers to how much people attribute the causes of events to internal or external factors. People with a high internal locus of control ("internals") tend to believe that most things that happen are their own fault, regardless of objective cause. On the other hand, those with a high external locus of control ("externals") tend not to accept blame for anything, preferring instead to believe in environmental reasons, even if they have clearly instigated an event. This stems from research by Rotter (1966) into the effects of reinforcement on preceding behaviour. People who believe that the outcome of a situation is controlled by external forces are less likely to raise or lower expectancies for future reinforcement following success or failure than those who perceive the reinforcement to be dependent upon their own skill or efforts. That is, externals do not make the connection between behaviour and reward as strongly as internals do, and this can consequently influence the learning process. People may have specific beliefs about certain situations, and they also differ in global attitudes to life in general. Rotter (1966) provides support for the hypotheses that internals: are more alert to those aspects of the environment which provide useful information for their future behaviour; take steps to improve their environmental condition; place greater value on skill or achievement reinforcements and are generally more concerned with their ability, particularly their failures; and are resistive to subtle attempts to influence them. Locus of control research can therefore clearly relate to driving, in areas such as skill and accident involvement. As such, there have been a few studies specifically addressing the issue of locus of control in the driving context.

In an attempt to measure driving locus of control, Montag and Comrey (1987) developed a specific driving internality-externality (I-E) scale. They asserted that previous I-E scales had a more successful record when tailored to be more domain-specific, rather than using a general scale. Montag and Comrey (1987) found that driving internality and driving externality are actually two independent constructs, as opposed to one bipolar dimension. Therefore, the use of two separate scales to measure these constructs is justified. They also found that externality correlates positively with fatal accident involvement, consistent with previous research which suggests that those with an external locus of control exhibit a lack of caution. Individuals with an internal locus of control are more attentive, motivated, and adept at avoiding aversive situations; hence internality is negatively related to accident involvement.

In developing these scales, Montag and Comrey (1987) used postdictive validation, and did not investigate the predictive strength of the scales. Arthur and Doverspike (1992) attempted to predict accident involvement using the Montag and Comrey (1987) Driving Internality-Externality (MDIE) scales. However, they found that locus of control was generally not associated with accident involvement, the only significant result was a positive relationship between driving internality and not-at-fault accidents - a result contrary to Montag and Comrey's (1987) findings.

Other research on driving locus of control has found that internality has been positively associated with alertness (Lajunen and Summala, 1995) and self-bias in accident involvement (Holland, 1993); whilst externality has shown to be positively correlated with aggression and tenseness (Lajunen and Summala, 1995), susceptibility to an alcohol placebo (Breckenridge and Dodd, 1991), and negatively correlated with perceived skill (Lajunen and Summala, 1995).

5. Mental workload

The topic of mental workload has been well documented in general texts (e.g., Sanders and McCormick, 1993; Singleton, 1989) and in more specific papers (Gopher and Kimchi, 1989; Schlegel, 1993; Wickens and Kramer, 1985). A general consensus is evident that there is no universally accepted definition of mental workload, although an analogy is often drawn with physical load. There has been a growing concern to rectify this state of affairs, for a quantifiable definition of mental load would have vast implications for safety and reliability, particularly in providing guidelines for what constitutes underload or overload (Singleton, 1989). This interest has stemmed primarily from the aviation industry (Sanders and McCormick, 1993). As Singleton (1989) points out, technology has made problems of mental load predominant, thus a solution may only come from multidisciplinary collaboration. An illustration of the diversity of the concept is reflected in the following research.

Schlegel (1993) views mental workload as a multidimensional interaction of task and system demands, operator capabilities and effort, subjective performance criteria and operator training and experience. This leads to the analogy between stress (task demands) and strain (impact upon the individual; capacity to meet demand). Similarly, Wang (1990) advocates a multidimensional construct of workload related to skills, motivation and emotion.

Alternatively, Leplat (1978) attempts to identify some of the factors determining workload. These may be divided into two broad classes - the worker and the conditions of work (though Leplat states that the critical problems here are of interactions between the two). Thus the specific factors may result from the requirements of the task (overload and underload); anatomical/physiological factors (e.g., fatigue, which has a cause and effect relationship - limiting possibilities thus forcing the operator to choose less efficient measures); factors of the physical surroundings; psychological factors (skill, personality, motivation); social factors (rules/conditions of work); and factors outside work. In addition, Okada (1992) finds that mental workload decreases as the number of indefinite (variable) factors involved in decision-making decreases.

A more theoretical approach is offered by Gopher and Kimchi (1989). Two models are posited: a computational model (parallel distributed processing) and one of behavioural energetics (motivational and intensive aspects). The former of these emphasises processing, the latter is concerned with resources. By integrating these, workload is viewed as the balance of automatic (attention-free) and controlled (resource-demanding) processing.

Mental workload is a major concern for automation. Reinartz and Gruppe (1993) argue that automated systems present cognitive demands, which *increase* workload. In their view, operators and the automated system are members of the same team. Thus effective control is dependent upon how well that team works and communicates together. The performance of the operator is hindered by the increase in processing load resulting from the additional task of collecting information about the system state. This is further complicated by the extent of the operator's knowledge about the system. In the event of manual takeover, the operator must either disable interlocks to other systems, or else match his/her actions to those of related process functions.

These issues are generally symptomatic of the transition in the role of the human from operational to supervisory control (Parasuraman, 1987). Such a situation has the paradoxical potential for imposing both overload and underload: reduced attention during normal operations, however difficulties increase when faced with a crisis or system failure (Norman, 1990). In the latter scenario, the human is forced to immediately return to the operator role, gather information about the system state, make a diagnosis and attempt a resolution. It has been argued (e.g., Wilson and Rajan, 1995) that whilst physical workload should be minimised, mental workload needs to be *optimised* to prevent performance decrements. This has been demonstrated in a driving context by Matthews and Desmond (1997), who found that stress impaired performance in underload conditions, suggesting a breakdown of adaptive regulation of effort.

Such concerns about workload have prompted many authors to review the issue of task allocation. Goom (1996) states that determining the optimum workload for the human should be the clearest single driver for the allocation process. Others have espoused the merits of dynamic task allocation, which can improve performance and workload ratings under situations of high demand (Tatersall and Morgan, 1997), and even in more complex cognitive tasks (Hilburn, 1997). These extend to investigations of MWL in car driving. Certain driving tasks have been found to increase mental workload and consequently present potentially dangerous situations (see e.g., Dingus, Antin, Hulse and Wierwille, 1988; Hancock, Wulf, Thom and Fassnacht, 1990). Some researchers (e.g., Brookhuis, 1993; Fairclough, 1993; Wildervanck, Mulder and Michon, 1978) have therefore explored the use of monitoring systems to detect situations of driver underload or overload, and to intervene either directly or indirectly if the situation becomes critical. This type of monitoring popularly utilises physiological channels. Other researchers (e.g., Schlegel, 1993; Verwey, 1993) are interested in the determinants of driver workload with a view to developing adaptive interfaces which may reduce such workload in future. These studies are generally in recognition of the fact that modern drivers are presented with increasingly complex information, with the advent of Intelligent Vehicle Highway Systems (IVHS). These ideas have been applied to driving in the form of adaptive driver systems such as GIDS, which aim to improve the quality of system behaviour by designing a human-centred system (Hancock and Verwey, 1997). Future technologies such as ACC will increase this need radically.

6. Driver stress

Stress has been defined as that which we appraise as harmful, threatening, or challenging, and research into stress has received a great deal of attention in the latter half of this century (see e.g., Cox and Griffiths, 1995; Holmes and Rahe, 1967; Wortman and Loftus, 1992). This is probably due to the fact that more people have realised the important role which stress plays in physical and psychological health, and its consequent impact on performance. Wortman and Loftus (1992) review strong evidence of a connection between stress and lowered immune system function, although it is not clear whether this is due to a direct effect on the immune system or indirect results of depression, poorer health habits etc.

Wickens (1992) reviews the effects of stressors on performance and human error. Arousal has long been associated with performance, as demonstrated by the classic inverted-U function of the Yerkes-Dodson law (Yerkes and Dodson, 1908). This was later modified by Easterbrook (1959) into cue utilisation theory. This effectively states that excessive arousal can affect the selectivity of attention, such that there is a narrowing (or “tunnelling”) of attention to different environmental or internal cues. The phenomenon of attentional narrowing under stress has been applied to areas as diverse as recovered memories of childhood sexual abuse (see Memon and Young, 1997, for a review) and the design of peripheral vision displays (Stokes, Wickens and Kite, 1990).

Stressful events can be one-off occurrences that happen to the individual (acute stress or daily hassles), or lifestyle changes instigated by the individual (chronic stress). Such events need not necessarily have negative connotations - getting married, for example, can be a source of stress. One common theme to all stressful events is that they evoke some adaptation or coping response. Indeed, the role of appraisal is primary in stress management, and this involves two steps (Lazarus, 1990; cited in Wortman and Loftus, 1992). Primary appraisal refers to a judgement of whether the event or situation is a threat to our well-being, whereas secondary appraisal involves determining whether we have the resources to cope with the threat. Such appraisal can very much determine individual strategies of coping, which can also take two forms (Lazarus and Folkman, 1984; cited in Wortman and Loftus, 1992). Problem-directed coping is an attempt to do something constructive about the stressful situation, and emotion-focused coping is an effort to regulate the emotional consequences of a stressful situation.

Driving in high congestion increases state stress, and can elicit a whole range of coping behaviours. Whilst individuals may express a general preference for direct (problem-focused) coping, when faced with high congestion there emerges an equal split between direct and indirect (emotion-focused) coping behaviours (Hennessy and Wiesenthal, 1997). Moreover, stress has been linked with road traffic offending, whether the source of stress comes from driving or other aspects of life (Simon and Corbett, 1996).

A good deal of the research into the stress of driving has been conducted by Gerry Matthews and his colleagues (Gulian, Glendon, Matthews, Davies and Debney, 1990; Matthews and Desmond, 1995a, 1995b). Their research uses the Dundee Driving simulator to investigate stress reactions in a variety of situations. Gulian et al. (1990) found that stress is associated with time of day as well as driving conditions, in that elevated stress levels were observed in the evening and in midweek. More importantly, the results suggested that driving stress is a global process with multiple causes, both intrinsic and extrinsic to driving. By simulating loss of control on an icy road, Matthews and Desmond (1995b) reported that this source of stress increases task-related interference and decreases perceived control. There was also evidence that the combination of stress and fatigue can disrupt the adaptive regulation of effort, such that their detrimental effects are exacerbated in low workload conditions. A logical conclusion from these results is that stress is an influence on accident risk.

Matthews and Desmond (1995a) considered the role of stress in the design of in-car driving enhancement systems. It was concluded that any adverse effects of in-car guidance systems (e.g., overload/attentional distraction) may be accentuated when under stress. Again, the mechanism for this was thought to be a disruption of effort mobilisation to match the varying workload of the task.

7. Situation awareness

Situation awareness (SA) is defined as "...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endlsey, 1995, p. 36). It is a state of knowledge of a dynamic system, rather than a process, and is separate from decision-making and performance.

The concept of SA appears to have emerged from the anecdotes of aircrew in operational contexts. The term was first coined in research texts in the latter half of the 1980s. Since then it has primarily been studied by researchers in the defence and aviation industries, who popularly quote poor SA as a major source of human error.

A related notion is alluded to by Woods (1988), who writes about situation assessment in complex dynamic systems. This is more concerned with problem solving as the state of the world changes. For instance, consider the incident that occurred at the Three Mile Island nuclear power plant. An initiating event sets of a chain of failures such that the problem is continually evolving over time. The problem solver must therefore keep track of this development and constantly update their understanding of the situation. Otherwise, they might fixate on a snapshot of the system, and attempts to resolve the problem at that point in time may not be relevant further down the line. Adaptability is therefore the key, and this can go wrong if the problem solver either fails to adapt or adapts in an inappropriate manner. How closely this is related to defined SA is a matter of debate. There is certainly some overlap between SA and mental representations. We believe this is a debate worth pursuing, so it is returned to later.

Research in SA has primarily been instigated by Mica Endsley (Endsley, 1995; Endsley and Kiris, 1995; Jones and Endsley, 1996), who has developed a model of SA in dynamic systems. SA is hypothesised to consist of three levels. Level 1 SA is the perception of elements in the environment, level 2 SA is the comprehension of the current situation, and level 3 SA is the projection of the future status of the system. These levels translate onto a taxonomy of SA errors, which are particularly relevant in automated systems, where the out-of-the-loop performance problem can result from a loss of SA. Most SA errors arise from a failure to obtain the needed information, a problem that will only be exacerbated in automated systems. Since SA is a product of applying skills (MacLeod, 1997), any skill degradation caused by automation will undermine SA, which can be particularly malignant following automation failure.

Although SA is relevant in a variety of systems operation, it seems to be most popularly applied in aviation and air traffic control (ATC). SA is hypothetically related to the creation and maintenance of the air traffic controller’s ‘picture’, a vital concept in ATC research (Isaac, 1997).

8. Mental representations

The concept of a ‘picture’ of one’s working environment has a high degree of overlap with the theory of mental models, or mental representations. Johnson-Laird (1983, 1989) describes mental models as a dynamic representation or simulation of the world. Indeed, our experience and perception of events is indirect, depending on our ability to construct models of them: “...what we perceive depends on both what is in the world and what is in our heads” (Johnson-Laird, 1989; p. 471). These models can either be physical or conceptual, and can be constructed directly (via observation), vicariously (through explanation), or indirectly (by analogy). Mental models theory is intended to explain the higher processes of cognition, particularly comprehension and inference. Johnson-Laird (1989) goes on to describe how mental models can be applied to syllogistic reasoning or to the representation of knowledge and expertise. It is the latter case that is most relevant here, however some explanation of models in syllogistic reasoning is pertinent. Given a set of premises, a model is constructed from them and a putative conclusion formed. This conclusion is then subjected to falsification attempts, and if it is not possible to find any counterexamples to it, the inferred conclusion is judged to be valid. Most of this takes place without conscious awareness, and errors can arise if the conclusion calls for more than one possible model.

Whilst this review is primarily concerned with the application of models in the real world, it would be incomplete if it did not acknowledge the theoretical research into mental models. Rips (1986) explores the use of mental models as semantic models, as simulations, and compares models to inference rules. Rips has problems with the mental model concept, finding that in most realistic situations, probabilistic reasoning is more plausibly explained through rules of thumb than simulation. Meister (1990) is more forgiving, stating that models can be used as simulations to simplify planning and decision-making. Finally, Brewer (1987) attempts to clear up any confusion over terminology between schemas, mental models, and imagery. In brief, schemas are generic mental structures underlying knowledge and skill, whilst mental models are inferred representations of a specific state of affairs which give rise to images. According to Brewer, the models relevant to this review are strictly termed *causal mental models*. That is, a domain-specific subclass of mental models which use causal representations to deal with physical systems (cf. the simulations mentioned previously).

So we turn to mental models of the world. In attempting to understand physical systems, it is hypothesised that people develop approximate representations in their heads, which are incomplete, unstable, ad hoc and unscientific, even superstitious (Eysenck and Keane, 1990). As an example, consider how thermostat on a home heating system works (Kempton, 1986). There are generally two theories which naive people hold, a feedback theory and a valve theory. Although the valve theory is not technically correct, it does prove functional and can even lead to better heat management. Around 25-50% of people subscribe to the valve theory, due to the fact that these models are based on everyday experience. This also explains why around 75% of people hold naive physical laws about projectile motion, which do not concur with the fundamental laws. Intriguingly, these 'natural' laws which are abstracted from experience reflect early pre-Galilean physics.

In a more specific example, Bainbridge (1992) uses industrial process operation to investigate the role of mental models in cognitive skill. The interpretation is reminiscent of the role of models in syllogistic reasoning. System information (e.g., from an alarm panel) leads to multiple inferences about the system state based on the operator's knowledge or models. The operator then begins an active search for confirmatory evidence about the inferences s/he has made. Skill plays its part if there is no direct information on system state, then the operator assumes a model and controls the system on prospective anticipation, rather than retrospective feedback. It is therefore crucial that the assumed model is correct, and Bainbridge suggests that in practice, operators should be provided with supplementary overviews of the system to enhance their mental models. This is especially important in automated systems, when the overview may be less apparent from direct interaction. Indeed, Wilson and Rajan (1995) believe that the notion of mental models, whether accurate or inaccurate, is useful for ergonomists for precisely that reason:

> "...if we can predict or understand even in some fashion what mental models a new operator or user might hold about a system and its relevant domain, and what model they might build through subsequent interaction with the system, then we can improve interface design, training operating procedures and so on. By understanding the potential users' mental models, and by adapting their own conceptual model accordingly, designers might develop a 'system image' that better matches, sustains and helps develop an appropriate user mental model." (p. 373)

9. A Research Framework and Hypothesised Psychological Model

From the review of the literature, we believe that the interdependency between the psychological concepts underrepresented. The mental model that a user develops about a system is critical to performance and operations with that system. Given the previous review of psychological factors associated with the use of automation, there are some obvious (and some not-so-obvious) interrelations between the variables. From the literature, it would seem that mental workload plays a central role in the relationship. For instance, it is apparent that high workload in the form of traffic congestion can increase stress (Wilson and Rajan, 1995), but there is some evidence that this relationship is bi-directional. Matthews and Desmond (1995a; 1997) provide evidence for the mechanism behind this relationship, and from this there are two novel yet logical conclusions relating workload to stress. The first is that stress can affect performance in low as well as high workload conditions. The second is that the effort involved in coping with stress actually adds to the task demands.

The question of whether feedback affects mental workload is contentious. Becker, Warm, Dember and Hancock (1995) found that performance feedback generally lowered mental workload in a monitoring task. However, the results of Fairclough et al. (1997) suggest that time headway feedback has no effect on workload in a car-following scenario. Either way, any relationship here is obviously unidirectional, and the results of the present study will hopefully shed some light on it.

Workload has also been known to affect situation awareness (Endsley, 1995; Jones and Endsley, 1996), however there is no evidence that this relationship is reversible. High workload is detrimental to SA, as attentional resources are primarily engaged in maintaining performance rather that SA. Indeed, workload is evidently a causal factor in approximately 30% of SA errors.

Situation awareness is also related to some of the other factors in this review. Stress weakens level 1 SA (perception) by causing attentional narrowing (Endlsey, 1995). Trust can also affect SA, as Jones and Endlsey (1996) found that 2.7% of level 1

SA errors were due to overreliance on automation. Finally, as has been stated previously, SA is highly related to mental models theory. Endsley (1995) sees SA as a situation model, or a context-specific mental model. This is supported by evidence from Jones and Endsley (1996) that 6.9% of level 2 SA errors, and 0.4% of level 3 SA errors are due to poor mental models.

From all of this evidence, then, we can construct a hypothesised model of the relationship between these factors as shown in figure 2. It must be stressed that although the individual links have been established in the scientific literature, the assembly of the model is purely conjectural for now. However it is our intention to investigate this model and, using structural equation modeling, to refine it and determine exactly what influences driver behaviour with automation.

FIGURE TWO ABOUT HERE

Figure 2. Hypothesised relationship between psychological factors

10. Conclusions

We have proposed a psychological model of the driver when using vehicle automation, which we believe has distinct advantages for applied research. The strength of the model lies in the fact that it is derived from the established literature in human factors and psychology. Furthermore, the literature review covered a variety of domains, from automotive, through human supervisory control, to aviation. All of the direct links between elements in the model have been reported elsewhere, but this paper assembles all of the factors in a unified structure for the first time.

Moray (1999) argues that “we need to offer … strong, preferably quantitative models of human-machine interaction” (p. 233). He describes three main advantages of modeling in research. Firstly, since engineers are used to working with quantitative models of the physical processes, presenting psychological processes by the same method can help ergonomists to communicate their arguments. Moray argues that the ergonomics discipline is now in a position to begin formulating such decent quantitative models. Next, quantitative models that are supported by data enable researchers to make predictions about behaviour in any given situation. The model we have proposed has been derived from a broad knowledge base, which we then intend to test in a specific environment, namely vehicle automation. Following the validation and refinement of the model, it is hoped that we can then feed back to make more generic quantitative predictions, for other automotive situations as well as other domains. This brings us on to the final advantage for modeling, which is to guide the design of automated systems. Currently, there is little in the way of design guidance from a theoretical base for automotive applications. The few studies in this area have been comparative studies, using traditional hypothesis testing to make their point. However, Moray states that "hypothesis testing does not lead directly to design recommendations in the way that modeling does" (p. 230). In particular, he argues that cognitive ergonomics should reject hypothesis testing in favour of the modeling approach to advance research and development. We hope that the model proposed in this paper represents a step in that direction.

In conclusion, then, we have attempted to bring together all of the psychological factors involved with using automation into one single model. Such a meta-analysis has been based on a considerable research effort using traditional hypothesis testing methods. Now, though, rather than examine one or two factors in isolation, it is necessary to test the interplay of factors within this model if useful design recommendations are the ultimate goal.

References

ANNETT, J., and KAY, H. 1957, Knowledge of results and 'skilled performance'. *Occupational Psychology*, **31**, 69-79.

ARTHUR JR, W., AND DOVERSPIKE, D. 1992, Locus of control and auditory selective attention as predictors of driving accident involvement: a comparative longitudinal investigation. *Journal of Safety Research*, **23**, 73-80.

BAINBRIDGE, L. 1992, Mental models in cognitive skill: the example of industrial process operation, in Y. Rogers, A Rutherford, and P.A. Bibby (eds*), Models in the mind: theory, perspective and application* (London: Academic Press), 119-143.

BARBER, B. 1983, *The Logic and Limits of Trust.* (New Brunswick, NJ: Rutgers University Press).

BECKER, A. B., WARM, J.S., DEMBER, W.N., and HANCOCK, P.A. 1995, Effects of jet engine noise and performance feedback on perceived workload in a monitoring task. *International Journal of Aviation Psychology*, **5**, 49-62

BILLINGS, C. E. 1997, *Aviation Automation: the search for a human-centred approach.* (Mahwah, NJ: Lawrence Erlbaum Associates)

BONNER, J.V.H. 1998, Towards consumer product interface design guidelines, in N.A. Stanton (ed.), *Human factors in consumer products* (London: Taylor and Francis), 239-258.

BRECKENRIDGE, R. L., and DODD, M.O. 1991, Locus of control and alcohol placebo effects on performance in a driving simulator. *Perceptual and Motor Skills* **72**, 751-756.

BREWER, W.F. 1987, Schemas versus mental models in human memory, in: P. Morris (ed.), *Modelling cognition.* (Chichester: Wiley), 187-197.

BROOKHUIS, K.A. 1993, The use of physiological measures to validate driver monitoring, in A. M. Parkes and S. Franzen (eds.) *Driving future vehicles*. (London: Taylor and Francis), 365-376.

COX, T., and GRIFFITHS, A. 1995, The nature and measurement of work stress: theory and practice, in J.R. Wilson and E.N. Corlett (eds.), *Evaluation of human work: a practical ergonomics methodology*. London: Taylor and Francis), 783-803.

DINGUS, T.A., ANTIN, J.F., HULSE, M.C., and WIERWILLE, W.W. 1988 Human factors issues associated with in-car navigation system usage. (An overview of two in-car experimental studies). In Anonymous. Proceedings of the Human Factors Society 32nd annual meeting (Santa Monica: HFES), 1448-1452.

DUNCAN, J., WILLIAMS, P., and BROWN, I. 1991, Components of driving skill: experience does not mean expertise. *Ergonomics*, **34**, 919-937.

EASTERBROOK, J. A. 1959, The effect of emotion on cue utilisation and the organisation of behaviour. *Psychological Review*, **66**, 183-201

ENDSLEY, M.R. 1995, Toward a theory of situation awareness in dynamic systems. *Human Factors*, **37**, 32-64

ENDSLEY, M.R., and KIRIS, E.O. 1995, The out-of-the-loop performance problem and level of control in automation. *Human Factors*, **37**, 381-394

EYSENCK, M. W., and KEANE, M.T. 1990, *Cognitive psychology: as student's handbook.* (Hove: Lawrence Eribaum Associates).

FAIRCLOUGH, S. 1993, Psychophysiological measures of workload and stress, in A.M. Parkes and S. Franzen (eds.), *Driving future vehicles*. (London: Taylor and Francis), 377-390.

FAIRCLOUGH, S. H., MAY, A.J., and CARTER, C. 1997, The effect of time headway feedback on following behaviour. Accident Analysis and Prevention, **29**, 387-397.

GOOM, M. 1996, Real world constraints on the allocation of functions, in S. A. Robertson (ed.), *Contemporary ergonomics 1996*, (London: Taylor and Francis), 300-305.

GOPHER, D., and KIMCHI, R. 1989, Engineering psychology. *Annual Review of Psychology*, **40**, 431-455.

GROEGER, J. A. 1997, *Memory and remembering: everyday memory in context*. (Harlow: Longman).

GULIAN, E., GLENDON, A.I., MATTHEWS, G., DAVIES, D.R., and DEBNEY, L.M. 1990, The stress of driving: a diary study: *Work and Stress*, **4**, 7-16.

HANCOCK, P.A., and VERWEY, W.B. 1997, Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, **29**, 495-506.

HANCOCK, P.A., WULF, G., THOM, D., and FASSNACHT, P. 1990, Driver workload during differing driving manoeuvres. *Accident analysis and prevention*, **22**, 281-290.

HENNESSY, D.A., AND WIESENTHAL, D. L. 1997, The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics*, **40**, 348-361.

HILBURN, B. 1997, Dynamic decision aiding: the impact of adaptive automation on mental workload, in D. Harris (ed.), *Engineering psychology and cognitive ergonomics.* (Aldershot: Ashgate), 193-200.

HOLLAND, C.A. 1993, Self-bias in older drivers' judgements of accident likelihood. *Accident Analysis and Prevention*, **25**, 431-441.

HOLMES, T.H., and RAHE, R. H. 1967, The Social Readjustment Rating Scale. *Journal of Psychomatic Research*, **11**, 213-218.

ISAAC, A. R. 1997, Situational awareness in air traffic control: human cognition and advanced technology, in D. Harris (ed.) *Engineering psychology and cognitive ergonomics.* (Aldershot: Ashgate), 185-191.

JOHNSON-LAIRD, P.N. 1983, *Mental models: towards a cognitive science of language, influence and consciousness.* (Cambridge: Cambridge University Press).

JOHNSON-LAIRD, P.N. 1989, Mental models, in M.I. Posner (ed.), *Foundations of cognitive science* (Cambridge, MA. MIT Press), 469-499.

JONES, D.G., and ENDSLEY, M.R. 1996, Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine*, **67**, 507-512.

KABER, D.B., and ENDSLEY, M.R. 1997, Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress*, **16**, 126-131

KEMPTON, W. 1986, Two theories of home heat control. *Cognitive Science*, **10**, 75-90.

KLUGER, A.N., and ADLER, S. 1993, Person versus computer-mediated feedback. *Computers in Human Behaviour*, **9**, 1-16.

LAJUNEN, T., l. and SUMMALA, H. 1995, Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. *Personality and Individual Differences*, **19**, 307 – 318.

LEE, J. D. and MORAY, N. 1994, Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, **40**, 153-184.

LEPLAT, J. 1978 Factors determining work-load. *Ergonomics*, **21**, 143-149.

MACLEOD, I.S. 1997, System operating skills, cognitive functions and situational awareness, in D. Harris (ed.), *Engineering psychology and cognitive ergonomics*. (Aldershot: Ashgate), 299-306.

MATTHEWS, G., and DESMOND, P.A. 1995a, Stress as a factor in the design of in-car driving enhancement systems. *Le Travail Humain*, **58**, 109-129.

MATTHEWS, G., and DESMOND, P.A. 1995b, Stress reactions in simulated driving. Symposium on the design and validation of driving simulators Valencia, Spain.

MATTHEWS, G., and DESMOND, P.A. 1997, Underload and performance impairment: evidence from studies of stress and simulated driving, in D. Harris (ed.), *Engineering psychology and cognitive ergonomics.* (Aldershot: Ashgate), 355-361.

MCFADDEN, S.M., GIESBRECHT, B.L., and GULA, C.A. 1998, Use of an automatic tracker as a function of its reliability. *Ergonomics*, **41**, 512-536.

MEISTER, D. 1990, Simulation and modeling, in J. R. Wilson and E. N. Corlett (eds), *Evaluation of human work: a practical ergonomics methodology*, (London: Taylor and Francis), 202-228.

Memon, A., and Young, M. 1997 Desperately seeking evidence: the recovered memory debate. *Legal and Criminological Psychology*, **2**, 131-154.

MONTAG, I., and COMREY, A. L. 1987, Internality and externality as correlates of involvement in fatal driving accidents. *Journal of Applied Psychology*, **72**, 339-343.

MORAY, N. 1999, The psychodynamics of human-machine interaction, in D. Harris (ed.), *Engineering psychology and cognitive ergonomics: Vol. 4* (Aldershot: Ashgate), 225-235.

MUIR, B.M. 1994, Trust in automation: Part 1. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, **37**, 1905-1922.

MUIR, B. M. and MORAY, N. 1996, Trust in automation. Part 2. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, **39**, 429-460.

NORMAN, D.A. 1990, The ‘ problem’ with automation: inappropriate feedback and interaction, not ‘over-automation’. *Phil. Trans. R. Soc. Lond. B.* **327**, 585-593.

OKADA, Y. 1992, Human characteristics at multi-variable control when the operator changes over the plant-lines. *Ergonomics*, **35**, 513-523.

PARASURAMAN, R. 1987, Human-computer monitoring. *Human Factors*, **29**, 695-706.

PARASURAMAN, R. 1997, Humans and automation: use, misuse, disuse, abuse. *Human Factors*, **39**, 230-253.

PARASURAMAN, S., SINGH, I.L., MOLLOY, R., and PARASURAMAN, R. 1992, Automation-related complacency: a source of vulnerability in contemporary organisations. *IFIP Transactions A Computer Science and Technology*, **13**, 426-432.

REINARTZ, S.J., and GRUPPE, T.R. 1993, Information requirements to support operator-automatic co-operation. *Human Factors in Nuclear Safety Conference,* (London, April 22-23 1994).

RICHARDSON, M., BARBER, P., KING, P., HOARE, E., and COOPER, D. 1997, Longitudinal driver support systems. *Proceedings of Autotech ’97* (London: ImechE), 87-97.

RIPS, L.J. 1986, Mental muddles, in M. Brand and R. M. Harnish (eds.), *The representation of knowledge and belief. (*Tucson: University of Arizona Press), 258-286.

ROTTER, J.B. 1966, Generalised expectancies for internal versus external control of reinforcement. *Psychological Monographs*, **80**, Whole no. 609.

SANDERS, M.S., and McCORMICK, E.J. 1993, *Human factors in engineering and design*. (New York: McGraw-Hill).

SCHLEGEL, R.E. 1993, Driver mental workload, in B. Peacock and W. Karwowski (eds.), *Automotive ergonomics*. (London: Taylor and Francis), 359-382.

SIMON, F. and CORBETT, C. 1996, Road traffic offending, stress, age and accident history among male and female drivers. *Ergonomics*, **39**, 757-780.

SINGLETON, W.T. 1989, *The mind at work: psychological ergonomics*. (Cambridge: Cambridge University Press).

STANTON, N.A., and MARSDEN, P. 1996, From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. *Safety Science*, **24**, 35-49.

STANTON, N. A. and YOUNG, M. S 1998, Vehicle automation and driver performance *Ergonomics*, **41**, 1014-1028.

STANTON, N. A., YOUNG, M. S. and McCAULDER, B. 1996 Drive-by-wire: the case of driver workload and reclaiming control with Adaptive Cruise Control. *Safety Science*, **27**, 149-159.

STOKES, A.F., WICKENS, D.D., and KITE, K. 1990 *Display technology: human factors concepts.* (Warrendale, PA: Society of Automotive Engineers Inc.).

TATERSALL, A.J., and MORGAN, C.A. 1997, The function and effectiveness of dynamic task allocation, in D. Harris (ed.), Engineering psychology and cognitive ergonomics. (Aldershot: Ashgate), 247-255.

TUCKER, P., MACDONALD, I., SYTNIK, N.I., OWENS, D.S., and FOLKARD, S. 1997, Levels of control in the extended performance of a monotonous task, in S. A. Robertson (ed.), *Contemporary ergonomics 1997*. (London: Taylor and Francis), 357-362.

VERWEY, W.B. 1993, How can we prevent overload of the driver? in A. M. Parkes and S. Franzen (eds), Driving future vehicles, (London: Taylor and Francis), 235-244.

WANG, Z. 1990, Recent developments in ergonomics in China. *Ergonomics*, **33**, 853-865.

WELFORD, A.T. 1968, *Fundamentals of skill*. (London: Methuen)

WHITE, J., SELCON, S.J., EVANS, A., PARKER, C., AND NEWMAN, J. 1997, An evaluation of feedback requirements and cursor designs for virtual controls, in D. Harris (Ed.) *Engineering psychology and cognitive ergonomics*. (Aldershot: Ashgate), 65-71.

Wickens, C.D. 1992, *Engineering psychology and human performance (2nd edition).* (New York: Harper and Collins).

Wickens, C.D., and Kramer, A. 1985, Engineering psychology. *Annual Review of Psychology*, **36**, 307-348.

Wildervanck, C., Mulder, G., and Michon, J.A. 1978, Mapping mental load in car driving. *Ergonomics*, **21**, 225-229.

Wilson, J.R., and Rajan, J.A. 1995, Human-machine interfaces for systems control, in J.R. Wilson and E.N. Corlett (eds), *Evaluation of human work: a practical ergonomics methodology*. (London: Taylor and Francis), 357-405.

Woods, D.D. 1988 Coping with complexity: the psychology of human behaviour in complex systems, in L.P. Goodstein, H.B. Anderson, and S.E. Olson (eds.), Tasks, errors and mental models: a festschrift to celebrate the 60th birthday of Professor Jens Rasmussen (London: Taylor and Francis), 128-148.

Wortman, C.B., and Loftus, E.F. 1992 *Psychology (4th ed.)* (New York: McGraw-Hill).

Yerkes, R.M. and Dodson, J.D. 1908, The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, **18**, 459-482.

Young, MS., and Stanton, N.A. (1997). Automotive automation: investigating the impact on drivers' mental workload. *International Journal of Cognitive Ergonomics*, **1**, 325-336.

Neville A. Stanton is a Professor of Human-Centred Design in the Department of Design at Brunel University, UK. He received his B.Sc. from the University of Hull and his M.Phil and Ph.D for the University of Aston in Birmingham, UK. His research interests span most psychological aspects of human interaction with technological systems. He has published his research in six books and over 50 journal papers. Professor Stanton is on the editorial boards of Theoretical Issues in Ergonomics Science, the International Journal of Human Computer Interaction, and Ergonomics.

Mark S. Young is a Research Fellow in the Department of Design at Brunel University, UK. He received a B.Sc. and Ph.D. from the University of Southampton, both in the Department of Psychology. His doctoral research was very much based in cognitive ergonomics, though, and it is this work he is carrying on at Brunel. His specific research interests concern the effects of advanced technology and automation on attention and mental workload.