Netherlands
organization for
applied scientific
research

IZF

TNO Institute for Perception

TDCK
Rapportencentrale ... 166394/6395
Frederik...
's-Gravenhage

Netherlands
organization for
applied scientific
research

TNO Institute for Perception

P.O. Box 23
3769 ZG Soesterberg
Kampweg 5
Soesterberg, The Netherlands

Phone +31 34 63 62 11

TNO-report

T.D. ...

1IZF 1990 B-3

W.B. Verwey

ADAPTABLE DRIVER-CAR INTERFACING AND MENTAL WORKLOAD: A REVIEW OF THE LITERATURE

13

The work discussed in this report is part of the project "Generic Intelligent Driver Support Systems" (GIDS) carried out under contract DRIVE V1041 of the European Community, in which the Traffic Research Centre, University of Groningen acts as prime contractor and Delft University of Technology, INRETS-LEN, Philips Research Laboratories, Saab Scania, TNO Institute for Perception, YA[illegible] Ltd., MRC-Applied Psychology Unit, Tregie Groupe Renault, VT[illegible] Universität der Bundeswehr, and University College Dublin are represented as partners. The opinions, findings, and conclusions expressed in this report are those to the author alone and do not necessarily reflect those of the EC or any other organization involved in the project.

Number of pages: 40

CONTENTS

Report No.:	IZF 1990 B-3
Title:	Adaptable driver-car interfacing and mental workload: A review of the literature
Author:	Drs.ing. W.B. Verwey
Institute:	TNO Institute for Perception
Date:	January 1990
HDO Assignment No.:	B89-59
No. in Program of Work:	735.4

SUMMARY

This report discusses a number of issues that are important for the design of an adaptive interface or dialogue controller in a car equipped with a GIDS system. There are two major sections, each followed by recommendations for interface design.
In the first section (Chapter 2), a short review of human performance in multitask situations is given. Practice is assumed to be of major significance for mental workload reduction especially in tasks that are executed concurrently with similar tasks. Also, issues concerning the concept, prediction and assessment of mental workload are discussed. A distinction is made between mental workload assessment for the evaluation of GIDS systems and for the adaptation of the interface to the driver workload. Recommendations are given for workload assessment in evaluation studies.
The second section of the report (Chapter 3) discusses literature on adaptive interfaces. Distinctions are made between human and system-initiated interface adaptation and between short-term and long-term adaptation. The section ends with recommendations regarding when and how to adapt the interface characteristics to driver workload.
Finally, the report concludes with a summary and issues for further research in chapter 4.

Rap.nr. IZF 1990 B-3 Instituut voor Zintuigfysiologie TNO, Soesterberg

Adaptieve bestuurder-auto interfaces en mentale werkbelasting: een literatuurstudie

W.B. Verwey

SAMENVATTING

Dit rapport behandelt een aantal onderwerpen welke van belang zijn voor het ontwerpen van een adaptieve interface, of dialoog controller, in een auto welke uitgevoerd is met een GIDS systeem. Het rapport is opgedeeld in twee secties, elk afgesloten door aanbevelingen voor interface ontwerp.

In de eerste sectie (hoofdstuk 2) wordt een overzicht gegeven van de kennis over menselijk functioneren in een multitaak omgeving. Oefening wordt aangenomen een belangrijke rol te spelen bij de reductie van mentale werkbelasting, met name indien taken samen uitgevoerd dienen te worden met gelijksoortige taken. Ook worden zaken behandeld die te maken hebben met het concept, de voorspelling en de meting van mentale werkbelasting. Een onderscheid wordt gemaakt tussen de meting van werkbelasting met als doel experimentele evaluatie van (GIDS) systemen en die met als doel de aanpassing van de interface aan de belasting van de bestuurder. Vervolgens worden aanbevelingen gegeven met betrekking tot de meting van werkbelasting in evaluatiestudies.

De tweede sectie van dit rapport (hoofdstuk 3) behandelt adaptieve interfaces. Er wordt een onderscheid gemaakt tussen interface adaptatie welke door het systeem en welke door de menselijke gebruiker wordt geïnitieerd. Daarnaast wordt onderscheiden tussen korte en lange termijn adaptatie. Deze sectie eindigt met aanbevelingen wanneer en hoe de interface-karakteristieken aangepast dienen te worden aan de werkbelasting van de bestuurder.

Het rapport besluit met een samenvatting en aandachtspunten voor verder onderzoek in hoofdstuk 4.

1 INTRODUCTION

The introduction of low-cost, high speed computers in traffic has just started. Anti-lock braking systems (ABS) and in-vehicle route guidance systems seem just some precursors of much more complex systems. Momentary thinking about future traffic electronics includes intelligent traffic signalling, traffic information systems, driver warning and assistance systems, automatic steering control, obstacle avoidance etc. (e.g. Shladover, 1988). Yet, despite enthusiastic generation of ideas and preliminary prototypes care should be taken for preventing the regular driver to be snowed under a myriad of flashing warning lights, demanding speech instructions, vibrating pedals, automatically turning steering wheels and flickering advices of which the poor regular driver does not understand a bit. After all, the driver is just trying to drive to her or his destination. Obviously, this is merely an exaggerated sketch of what should not be. Large scale introduction of electronics into traffic can indeed result in an enormous increase in road capacity, safety and driver comfort as stated by advocates but care should be taken that the design carefully matches the driver's capabilities.

The present paper is a review of the literature that is relevant for the design and evaluation of intelligent driver support systems. The review was carried out for the GIDS project which is supported by DRIVE. DRIVE is a research programme that has been initiated by the Commission of European Communities (CEC). The acronym DRIVE stands for Dedicated Road Infrastructure for Vehicle safety in Europe. The general aim of DRIVE is supporting background research to guide the introduction of new technologies in traffic. The formal objectives of the programme are defined as the improvement of road safety, transport efficiency and environmental quality (Commission of the European Communities, 1989). Research funded by the programme has started early 1989 and will last three years. The GIDS project (Generic Intelligent Driver Support) was one of the research proposals and was accepted under Project number V1041. The overall goal of the GIDS project is to determine requirements and design standards for a class of intelligent co-driver systems which will be maximally consistent with information requirements and performance capabilities of the human driver (Smiley and Michon, 1989). One of the major characteristics of GIDS systems will be that these systems adapt to the momentary workload of the driver, which is also a major issue in this report. In chapter two a review will be given of theoretical and more practical (i.e. mental workload) considerations of behaviour in multitask situations in order to come to recommendations about interface design and mental workload assessment in experimental situations. In the third chapter, literature on intelligent interfaces will be reviewed and recommendations concerning when and how to adapt, will be presented. Finally, in the last chapter recommendations for in-vehicle interfaces will be summarized and areas for further research will be indicated.

2 MULTIPLE TASK PERFORMANCE

One of the major questions about the design of a driver-car interface is how driver mental workload can be alleviated in order to prevent - at the least - a reduction of traffic safety. Theoretical research has prevented the usage of the term and preferred to utilize the term attentional requirements or, when speaking about multitask performance, divided attention. The theoretical literature is worth reviewing since it allows derivation of general recommendations for car-driver interfaces. By contrast, workload assessment is an atheoretical method to determine attentional requirements without giving cues how to reduce workload in a particular situation [with the multiple resource theory (Wickens, 1980, 1984) as the only exception]. So, in this chapter human performance in multitask environments will be reviewed from a more theoretical point of view. Next, the concept of workload and its assessment and use in a traffic environment will be considered. This will finally result in recommendations for the design and evaluation of GIDS-like driver-car interfaces.

2.1 Limitations in human information processing capacity

Early models of divided attention in experimental psychology emphasized that humans have a ***single limited capacity*** of information processing. Thus, in a multitask situation tasks were assumed to compete for a common processing structure (Broadbent, 1958) or resource (Moray, 1967) and humans would be unable to perform more than one 'thing' at the same time. That is, humans constitute a single-channel processor. The bottle-neck in human information processing was assumed to be located either at the perceptual level (Broadbent, 1958) or at the response level (Deutsch and Deutsch, 1963). The fact that humans are seemingly able to perform more tasks at the same time and thus divide their attention efficiently (e.g. keeping course while changing gears and maybe even talking at the same time) was explained by covert attention switching between the tasks (Broadbent, 1958, 1982). However, limited capacity theories have come under challenge in recent years as a result of several studies claiming to demonstrate little or no decrement in dual-task situations (e.g. Allport, Antonis, and Reynolds, 1972; Hirst, Spelke, Reaves, Caharak, & Neisser, 1980; Shaffer, 1975).

Multi-capacity theories of human information processing argue that human information processing depends on separate resources and that humans can divide their attention efficiently between concurrent tasks in case the tasks draw on separate rather than common processors or processing resources (McLeod, 1977; Wickens, 1980). The most influential of these multi-capacity theories is the multiple resource theory (Wickens, 1980, 1984a,b) which claims that tasks can be well executed concurrently when the tasks

1 utilize different modalities of input (visual versus auditory) and response (manual versus vocal),
2 differ in the demands on certain stages of processing (perceptual, central, or motor processes), and
3 demand different codes of perceptual and central processing (spatial versus verbal codes).

Later, Wickens (1984a; Wickens, Sandry & Vidulich, 1983) extended his theory with the stimulus-central processing-response (S-C-R) compatibility principle which asserts that spatial tasks are best accommodated by visual input and manual output whereas verbal tasks are best served by auditory input and speech output. Because of their empirical success in numerous applications, multiple resource theory and S-C-R compatibility have gained widespread acceptance. Yet, later research indicated shortcomings. Little evidence was found for resources in the input modalities when two discrete tasks are timeshared (Wickens, Fracker, and Webb, 1987) or no eye-movements are required (Wickens and Liu, 1988). Also, when an auditory and a visual task are being timeshared, the auditory task tends to attract full attention on the cost of the visual task (i.e. preemption; Wickens, Fracker, and Webb, 1987). Finally, multiple resource theory does not make detailed statements about the effects of practice.

2.2 Mental workload

The term 'mental workload' has been used several times in the previous sections without much clarification. This section ' ll mention some definitions and elaborate more on what is known about workload, assessment methods and the role it plays in traffic safety. The importance of workload is already intuitively apparent but has also been shown by findings that mental overload while driving a specific route covaried with the probability of traffic accidents on that route (MacDonald, 1979; Taylor, 1964). In the present context, workload is important in two ways. First, it is the tool that is used to evaluate GIDS systems in an experimental context. This indicates which system yields lowest driver workload and it leads to recommendations for the design of driver-car interfaces. Second, intelligent interfaces may automatically adapt to the momentary driver workload in a regular GIDS-system, but then the system requires access to some measure of driver workload. Obviously, these two applications of workload require different assessment methods and will therefore be considered separately. This chapter will address mental workload assessment in experimental conditions whereas on-line workload assessment by GIDS systems will be addressed in chapter 3.

The main reason for measuring mental workload has been the wish to predict performance of a man-machine system before it gets into production, to decide about task alterations, and to assess individual differences in task performance (Wickens, 1984a). According to Sanders (1979) the concept of mental workload, or mental load, is defined in common sense terms and (therefore) not consistently dealt with in models of human performance. Examples of workload definitions are "Workload is the extent to which an operator is occupied by a task" (Jahns, 1973), "... the term workload will be used to refer to the integrated effects on the human operator of task-related, situation-related, and operator-related factors that occur during the performance of a task" (Hart, 1986), and " ... the rate at which information is processed by the human operator, and basically the rate at which decisions are made and the difficulty of making the decisions" (Moray, 1979). Indeed, these definitions are rather abstract and not very clarifying. Probably this is caused by the fact that usage of mental workload measures have been of more concern to human factors practitioners than to theoretical psychologists and thus from the beginning emphasis was on how and when to measure rather than on theoretical and empirical foundation of the assessment methodology.

There are at least two basic problems with the concept of workload (Sanders, 1979). ***First***, when speaking about mental workload different researchers have different things in mind (Williges and Wierwille, 1979). Some refer merely to perceptual-motor workload when speaking about workload. Others incorporate the task environment with its physical, social, and emotional components. Then, the concept becomes almost equivalent to stress or strain. Frequently, though, the term workload is used to indicate load on the perceptual, central, and output resources (Wickens, 1980, 1984a). Then, 'mental' does not refer to workload on central resources but incorporates also workload on the in- and output modalities. The ***second*** problem is the absence of a standardized measure. Mental workload has been operationalized in a number of ways yielding workload scores that often correlate less than moderately. Taken the absence of any theoretical foundation of workload assessment (Colle, Amell, Ewry, and Jenkins, 1988) one measure is as good as another and it is difficult to equate different kinds of workload in order to compare one situation with another (Hart and Bortolussi, 1984).

2.3 Practice effects

With practice, mental workload in a complex task environment, such as driving, is largely reduced. Important mechanisms of learning in a multitask situation are the development of descriptive and procedural information (e.g. Anderson, 1983). Initially, one will obtain a mental model of the task environment. This model consists of a description of past experiences, stored in memory, which may serve to expect stimuli and allows for response preparation. This has been termed ***descriptive*** information (Anderson, 1983). ***Procedural*** information refers to knowledge that indicates what to do in a particular situation, and how. Procedural knowledge allows behaviour to become efficient since actions do not have to be selected from a large set any more, instead they are activated automatically (e.g. Logan, 1985). One may not even be aware of the presence of procedural knowledge (e.g. Willingham, Nissen and Bullemer, 1989). For example, a highly skilled sportsman may be able to carry out a practised sequence of actions without exactly knowing how he does it. So, procedural information requires extensive practice but enables one to perform efficiently without much attentional requirements (i.e. workload is low). This implies that tasks that result from a GIDS system should enable the driver to develop procedural knowledge in order to alleviate workload. The question then is how a GIDS system should look like for optimal development of procedural knowledge?

In general, skilled behaviour is characterized by efficient performance of individual and concurrent actions and efficient coupling of individual sequential actions. Efficient performance of individual actions has been labelled automatic behaviour (e.g. Schneider and Shiffrin, 1977) which, due to reduced attentional requirements, allows for increased task performance when the automatic task is time-shared with any other task (e.g. Schneider and Fisk, 1984). Tasks that are always executed simultaneously or in fixed sequence can integrate which also allows for efficient performance (Brown and Carr, 1989; Kahneman, 1973). Integrated tasks may utilize automatic processing for individual actions but this is not necessarily so (see below).

Automatic execution of individual actions develops as a function of consistent practice (Fisk, Ackerman, and Schneider, 1987; Schneider and Shiffrin, 1977) and makes explicit stimulus evaluation - and thus attention - unnecessary. Fisk, Ackerman, and Schneider (1987) have outlined six principles of human performance. Most important in the present context are:

1 Performance improvements will occur only for situations where stimuli can be dealt with the same way from trial to trial.

2 The human is limited, not by the number of mental operations he or she is required to perform, but by the number of inconsistent or novel task operations.

3 In order to alleviate high workload situations, consistent task components must be identified and, once identified, training of those components should be given in order to develop automatic component processes. Similarly, to make performance reliable under environmental stressors (such as fatigue, heat, stress, alcohol), training should be conducted to develop automatic task components.

4 For tasks requiring sustained attention (vigilance), automatic target detection should be developed by training prior to the execution of the task.

What should be clear from these principles is that the concept of consistency is intended as the cornerstone for any discussion concerning relevance of automatic processing theory to real-world skills. Examples of consistencies are utilizing the same stimuli (e.g. colour, letters, words, stimulus categories), the same relationships among stimuli and responses (e.g. switch up is always OFF and switch down is always ON), the same spatial locations for stimulus presentation and responding (e.g. visual navigation information is always presented on one particular location), and the same context (e.g. one kind of message refers only to one aspect in the environment). In addition, Gladstones, Regan, and Lee (1989) have shown that when there is no possibility to make task processing easier by automatization, performance will not improve much with practice - irrespective of multiple resource considerations like overlapping input and output modalities of concurrent tasks. This would indicate that when two or more genuinely independent and low-redundancy tasks require continuous attention, designers of complex systems should not assume much effect of practice and utilizing different in- and output modalities will not improve performance.

Besides automatization of individual actions, actions can also integrate into one molar unit with practice (Kahneman, 1973). Simultaneous actions can integrate because stimulus representations (Miller, 1956) and motor actions that initially belong to separate tasks integrate into one unit (e.g. Adams, 1984; Brown and Carr, 1989; Carswell and Wickens, 1987; Schneider and Detweiler, 1988; Wickens, Fracker, and Webb, 1985). Kahneman (1973) illustrated task integration with the example that when several simultaneously presented lights each require a response, subjects first start to respond to each light in a serial fashion. With extensive practice the appropriate responses are processed and executed simultaneously. Kahneman (1973) concluded that for each combination of lights one memory representation had developed and that each particular combination of lights triggers one particular combination of responses.

Actions may also integrate in case they are executed in rapid succession. An example of the integration of serial tasks in car driving is gear changing (Welford, 1988). Although not much is known about integration of sequential actions, Verwey (in preparation), in an initial attempt to tackle this issue, assumed that integration occurs when preparation for a forthcoming action occurs during a preceding action and/or when an action triggers a forthcoming action automatically. Such triggering may occur at the perceptual and the response level of processing. This model of integration was supported by experimental data. Although integration of tasks enables an enormous decrease in workload with practice, precise requirements for such integration of tasks are still unknown and further research is required.

So, practice is a major determinant of performance efficiency in task environments that require divided attention. The mo consistent the task environment, the greater the gains from practice. Practice may allow humans to reduce interference between separate tasks by forming a descriptive, mental model of the task which requires them to prepare upcoming actions and alleviates workload. Procedural knowledge incorporates individual actions that have automatized and/or integrated. As soon as combinations of initially separate tasks are integrated into one molar task, possibly with automatic properties, workload is even more decreased. Finally, when concurrently executed tasks are independent, low-redundant and require all attention, no efficient time-sharing strategies can be learned and separating input and output modalities does not improve performance or lowers mental workload.

2.4 Mental workload assessment

There are several ways to measure mental workload. These include primary task, secondary task, physiological, and subjective measures.

First, ***primary task*** performance can be used as indicator for workload under the assumption that an increment of task load reduces performance. Examples are error evaluation and reaction speed. Yet, when the human subject still has spare processing capacity left (i.e. he can handle more information without performance decrement), variations in workload will not show up in primary task measures when the difficulty of the primary task is increased. In that case attempts should be undertaken to establish the amount of spare processing capacity. Sanders (1979) proposed to increase the difficulty of the primary task until performance drops ("testing the limits"). Also, a secondary task may be utilized to assess spare processing capacity.

The ***secondary task*** technique, also referred to as the subsidiary task technique, assumes that an increase in workload on one task can be indicated by the level of performance on a second, low priority task. The main task has to be performed as good as possible and the workload of the task is indicated by performance on the secondary task. Hence, when the primary task gets more difficult less attention is available for the secondary task and performance drops. Yet, apart from the fact that the second task may well interfere with the primary task (e.g. Noy, 1987; Wierwille and Gutmann, 1978) the secondary task technique is often difficult to interpret as a measure of general workload because it assumes that humans are single capacity processors whereas empirical evidence has yielded ample evidence that humans possess several processing resources (see section 2.1). In other words, a secondary task will be a better indicator for workload of the primary task when it is more similar to the primary task. When a general picture of workload is required a battery of secondary tasks should be utilized (Kahneman, 1973). Still, the secondary task technique is well suited for assessment of the demands of one resource in a complex task that loads several resources. For example, when one wants to know whether a subject is visually loaded a secondary task can be adopted that requires much visual attention.

Note that with practice in a complex multitask environment one may change performance strategy. For instance, latency may be sacrificed for accuracy. Also, parts of a particular task may not be carried out any more or one task receives more attention. Therefore, it has become good practice to let the primary task be unaffected by the presence of the secondary task (e.g. by instruction or performance feedback), so that performance on the secondary task can be easily compared to performance of the secondary task in isolation. This emphasizes that instructions should be given not to drop performance on the primary task and several performance measures on all tasks should be evaluated in order to check whether task priorities or speed-accuracy trade-off has changed. So, the researcher should always be sure that an apparent variation in workload as concluded from secondary task performance is not a result of a change in the primary task performance (e.g. Noy, 1987).

Physiological measures can be applied to measure workload as well. Some measures are primarily sensitive to workload upon certain processing resources (e.g. event related potentials, e.g. P300, indicates perceptual/cognitive load) whereas others (e.g. pupil diameter and heart rate variability) seem to indicate overall mental workload (Kramer, Sirevaag, and Braune, 1987; Roscoe, 1987; Wickens, 1984a,b). Unfortunately, physiological measures are sensitive to artifacts like physical workload, noise and emotion-induced effects (Roscoe, 1987; Sanders, 1979). Also, physiological measures are fairly intrusive and they are sensitive to large inter-individual differences. Still, they are usually applicable without interfering with the task and, so, large efforts have been made to find physiological correlates of workload. From a large number of physiological indicators for workload one measure has in particular found to be promising for assessment of general workload. It has been proposed by Mulder (1973, 1979) and Sayers (1973). They noted that the 0.1 Hz component of the heart beat frequency in a frequency spectrum analysis (i.e. the heart-rate variability in the 0.1 Hz. region as indicated by a Fourier analysis) correlated with the amount of effort expanded by the operator in controlled attention demanding tasks (see also Aasman, Mulder and Mulder, 1987). This finding has been confirmed in several studies, e.g. Moray and his coworkers (Moray, Turksen, Aidie, Drascic, Eisen, Kruschelnicky, Money, Schonert, and Thornton, 1986; Vicente, Thornton, and Moray, 1987) who developed an on-line analysis technique.

Finally, advocates of ***subjective estimates*** or ratings of workload assert that humans are quite capable of estimating task difficulty. It is implicitly assumed that humans can judge average spare processing capacity over several processing resources. That subjective estimates are applicable in a wide range of task types has been concluded, among others, by Reid and his colleagues (e.g. 1985) and Hart and Staveland (1988). They studied subjective estimates in tasks including perceptual, memory, decision making and problem solving, and motor tasks. Especially when task demands are in the mental underload region, subjective estimates provide an accurate index of workload (Eggemeier and Stadler, 1984). Yet, when a task imposes demands in a such a manner that workload exceeds the capacity of working memory, subjective measures are less sensitive and performance measures should also be incorporated (Yeh and Wickens, 1988). Also, subjective estimates may not be obtainable during periods of very high workload due to the additional task demands of estimation in particular if these estimates involve several dimensions (Roscoe, 1987). Interestingly, some investigators (e.g. Ursin and Ursin, 1979) have suggested that physiological methods do not measure the imposed workload directly but instead yield information concerning subjective workload estimates and the ability to cope with it. When that is true then physiological measures of workload would be a continuous indicator of subjective workload estimates.

In addition to these workload assessment methods, there are also ways to ***predict mental workload*** from an analysis of the task and a - usually highly simplified - model of the human operator. The classical method is time-line analysis which assumes that workload is 100 percent when the operator is busy for 100 percent of the available time (Wickens, 1984a). Usually, time-line analysis is difficult since operator and task models are very complicated to make, do not assume effects of practice, consider the human solely as a single channel processor (Wickens, 1984a), do not consider effects of prior workload (Cumming and Croft, 1973), and do not reckon with effects of self-paced flexibility and strategies (Schneider and Detweiler, 1988). Clearly, taking the amount of time that the operator is not working is, indeed, a very crude measure of workload. However, recent developments in cognitive modelling (e.g. Anderson, 1983: ACT*; Laird, Newell and Rosenbloom, 1987: SOAR; Schneider and Detweiler, 1988: a connectionist/control architecture) seem very promising because these models are based on modern views of human information processing and incorporate effects of practice and strategies.

2.5 Which assessment method?

Given the large amount of workload measures, each with their specific pros and cons, the human factors practitioner is faced with the problem which measure, or combination of measures, to employ when she wants to evaluate workload of a specific system design. Nowadays, the general view is that workload is multidimensional (e.g. Roscoe, 1987; Wickens, 1984b) and that several measures should be selected in accordance with the dimensions of task load that are in greatest potential demand in the situation under evaluation (e.g. Bauer, Goldstein, and Stern, 1987; Roscoe, 1987). Thus, visual components of workload can, for instance, be evaluated using eye-movement registration. Perceptual and cognitive demands can be obtained by the P300 component in the EEG and response load can be evaluated by response observation or muscle activity assessment. In all these cases, however, an alternative is to utilize secondary tasks with the same presentation and response modalities, and central codes (i.e. the same resources according to the multiple resource theory, Wickens, 1984a,b) to determine spare processing capacity of each of the resources.

In the literature a number of studies have explicitly approached the question which assessment method is best in which situation. Williges and Wierwille (1979) reviewed 14 measures of workload and concluded that no one single measure can be recommended as the definitive behavioral measure of workload. Yet, they viewed subjective estimates and model-based workload prediction as the best candidates for workload estimation. In another paper reviewing several workload measures Johanssen, Moray, Pew, Rasmussen, Sanders and Wickens (1979) concluded that measures should be applied that are not task-specific. They proposed two measures that seemed interesting to pursue, the so-called 'difficulty margin' and the 'generalised Cooper-Harper scale'. The first is defined as the difference between the maximum possible performance over a considerable amount of time and the required performance. It, therefore, constitutes a primary task performance measure. The difference between maximum possible and r ,uired performance can be approached by increasing the difficulty of the task

(Sanders, 1979). The generalised Cooper-Harper scale consists of subjective estimates of workload[1]. Yeh and Wickens (1988) also urged the human factors practitioner to incorporate both performance measures (e.g. primary task performance) and subjective estimates in system evaluation because of the already mentioned dissociation between subjective and objective workload measures under highly demanding conditions. Finally, Roscoe (1987) established that in the realm of pilot workload subjective estimates in combination with heart rate is highly valuable to assess general workload.

2.6 Mental workload assessment in driving

It has become good practice to distinguish three levels in the driving task (e.g. Janssen, 1979). The control level is concerned with elementary vehicle handling functions like lane-keeping and handling of the controls. The manoeuvring level deals with aspects that have to do with car manoeuvring, overtaking, intersection negotiation and the like. Finally, the strategical level regards route planning and following. The levels differ with concern to the imposed load on different processing resources (Wickens, 1984a,b). Workload in navigation mainly loads memory and is therefore mainly central. Tasks at the control level of driving become largely automatic with practice. Therefore, these tasks do not load so much the central resources but rather the visual and motor resources. In general then, when one wants to establish workload in driving the appropriate assessment method depends largely on the kind of task the driver is performing.

Empirical research on the distinction between mental workload assessment methods in driving has also been done. Wetherell (1981) compared seven auditory-vocal tasks secondary to the driving task in order to find load on central resources. The subjects' task was to follow a quiet rural road (i.e. basically a control task). He found no one secondary task outstanding as a general measure of the workload imposed by driving. However, this is to be expected when one reasons from the view that central resources are barely loaded by car control. Still, in the light of recent theoretical work (e.g. Neumann, 1987; Schneider and Detweiler, 1988), Wetherell's (1981) conclusion that the

[1] Later, more sophisticated subjective workload assessment methods have been developed. The Subjective Workload Assessment Technique (SWAT) by Reid and coworkers (Reid, Shingledecker, and Eggemeier, 1981; Reid, Shingledecker, Nygren, and Eggemeier, 1981) and the NASA-TLX (Hart and Staveland, 1988) are well-known examples.

most appropriate secondary task is the one that is most similar to the additional task of interest, seems still sound.

Hicks and Wierwille (1979) compared five methods of measuring workload in a driving task in which gusts of wind at the front of the vehicle represented workload. In this task, which basically involved a control task too, subjective workload estimates and primary performance measures were most sensitive to workload. Visual occlusion, cardiac arrhythmia, and secondary task performance yielded no significant effects.

Gstalter (1985) measured subjective workload estimates, driver behaviour, heart rate, and galvanic skin response while driving through Munich, FRG. He concluded that workload indicators correlated rather low and that subjective estimates were most consistent and useful.

In a comparative study of mental workload assessment procedures in various driving conditions, Curry, Hieatt and Wilde (1975) employed four techniques, one psychophysiological (heart rate and heart rate variability) and three secondary task techniques (random digit generation, interval production and a short-term memory task). They found cardiovascular measures and interval production (e.g. Michon, 1966) to be most useful.

Biesta and Blaauw (1976) used heart rate, heart-rate variability and an auditory detection task to assess driver workload resulting from static and dynamic aspects of the environment. They found heart rate and especially heart-rate variability to be sensitive to static aspects of the environment (e.g. the kind of road). The auditory detection task, however, was found to be an insensitive measure for driver workload, whereas it did affect driving performance and the heart-rate measures. The authors suggested that subjects had given too much attention to the secondary task, not in the least because the stimuli were presented through headphones which made it hard to ignore them. These results can also be interpreted in the sense that auditory detection taps resources that are unoccupied in driving (Wickens, 1980, 1984a,b), which makes this secondary task unsuitable for assessing driver workload.

Finally, Janssen and Gaillard (1984) compared the 0.1 Hz component of the cardiac interval spectrum, the P300 evoked component of the EEG and driver urinary catecholamine excretion rates . a driver route choice experiment. They concluded that all three measures were satisfactory but that the 0.1 Hz cardiac interval spectrum component was somewhat more sensitive and consistent as measure of workload than both others.

In combination, these experiments suggest that subjective estimates and the 0.1 Hz cardiac component are a sensitive measure for general workload assessment in driving (cf. Roscoe, 1987). Assessment of workload on particular resources may well be done by various secondary tasks. However, it should be emphasized that utilizing the secondary task technique requires the primary task (here driving) to be unaffected by the presence of a secondary task. Noy (1987) presented a review of the driving research literature employing secondary task measures and found that some secondary tasks have adverse effects on at least some aspects of driving when they are performed concurrently. This emphasizes once more that the secondary task technique requires proper evaluation of performance in the driving task and any result where this has not been done would be suspect (Noy, 1987).

One final remark on driver workload concerns subject gender. In applying seven auditory-vocal secondary tasks, Wetherell (1981) found that performing the secondary task appeared to interfere with female driving ability, but not with that of males. In addition, Hancock (1988; 1989; Hancock, Rodenberg, Mathews, and Vercruyssen, 1988) showed that females rated subjective workload higher than their male counterparts. The robustness of these effects is not clear but, nevertheless, they suggest that for valid generalization gender should be incorporated as factor in driving studies where workload is assessed.

2.7 Recommendations

The review in this chapter allows derivation of recommendations for the design and evaluation of GIDS systems[2].

1 Multiple resource theory states that when tasks should be carried out concurrently, without possibilities to integrate, they should be different with respect to (1) input (visual versus auditory) and output (manual versus vocal) modalities, (2) perceptual/central and

[2]For more general guidelines concerning displays and controls the reader is referred to Galer and Simmonds (1984).

response processes, and (3) perceptual-central codes (verbal versus spatial) for optimal time-sharing. Since the control task of driving is mainly visual-spatial-manual, additional tasks should preferably be auditory-verbal. In case it is hard to make a task auditory-verbal, research should indicate what is better, a task that is mentally loading because it is presented and performed in an inefficient modality or a task that interferes more with driving because it requires the same resources (see for example Verwey, 1989). The S-C-R principle asserts that spatial tasks are best accommodated by visual input and manual output, whereas verbal tasks are best served by auditory input and speech output.

2 For general workload assessment in the evaluation of GIDS systems, subjective workload estimates (e.g. the SWAT) and the 0.1 Hz cardiovascular component are recommended. To evaluate dimensions of workload secondary tasks tapping specific processing resources, specific physiological measures (e.g. eye-movements, P300, muscle activity), or subjective estimates of each dimension seem better suited. When utilizing the secondary task technique, subjects should emphasize the primary task (usually driving). Measurement of the primary task should make sure that it has not deteriorated from the presence of the secondary task. Also, the particular secondary task should be similar to the task imposed upon the driver when using a GIDS system in order to measure spare processing capacity at the relevant resources. For example, workload resulting from navigation should be assessed by a method that is sensitive to central load (for example an auditory-vocal secondary task) and workload due to car control should be measured by assessing perceptual and motor load (e.g. a visual detection secondary task and a secondary task that utilizes manual responding).

3 GIDS systems should incorporate a possibility to adapt to the level of driving experience since inexperienced individuals work in a single channel mode, i.e. every task requires conscious attention and decision making and cause significant mental workload. Thus, novice drivers should not be loaded with additional tasks when not strictly necessary because they regard each task separately. Also, GIDS systems should be designed so that any driver task includes a high degree of consistency. This enables extended effects of practice. Consistency may, for example, imply using consistent colour coding over separate tasks (e.g. red always indicates danger), using similar buttons or switches for similar tasks (e.g. on is switch up) but different ones for incomparable tasks. Accord-

ing to recent views, practice integrates several tasks into one molar task. However, empirical evidence for task integration is still weak and effects of a variety of variables unknown. So, further research is required for more detailed recommendations on workload reduction with extended practice.

4 According to the preemption principle, auditory information is likely to attract the attention of the driver and should therefore be utilized with caution, especially when used with higher intensities (Sanders, 1983).

5 Finally, sex differences have been found in workload assessment while driving. For optimal generalizability it is, therefore, recommended to use both male and female subjects in the experimental evaluation of GIDS systems.

3 ADAPTIVE INTERFACING AND GIDS

One major feature of GIDS systems is that they possess an intelligent interface. Such interfaces should ***integrate*** several sources of information, ***adapt*** to the driver[3] workload by making things easier, ***know when to warn*** the driver and when not, and possibly ***take over*** parts of the task. In other words, an intelligent interface improves the efficiency of human-system interaction in keeping with the general human factors maxim "bend the tool, not the person" (McCormick and Sanders, 1982, Chapter 10). Rouse, Geddes, and Curry (1987) referred to the adaptable aspect of an intelligent interface as "an executive's assistant who zealously guards the superior's time and resources" (p.96), and according to Hancock, Chignell, and Loewenthal (1985) an adaptive interface would be like "a servo-mechanism that minimizes the difference between current demands and available operator capacity" (p.629). Thus, when a driver gets very busy with, for example, manoeuvring his car in city traffic the GIDS interface might adapt to the increased demands of driving by blocking phone calls and automatically turning the wipers on when necessary.

[3] In this chapter, frequent reference is made to literature that is not specific to the driving task. In case that literature speaks about the 'human operator' and the conclusions seem not directly applicable to driving the term 'human operator' is maintained in the report.

A distinction can be made between short-term and long-term adaptation. Short-term adaptation would be necessary because changes in the traffic environment directly affect driver workload. By contrast, long-term adaptation denotes that experienced drivers can use a more complex interface since they are less loaded by driving. So, long-term adaptation should not be sensitive to short-term variations in the environment.

When conceptualizing an adaptive interface two issues evolve. The first issue concerns the ***initiative of adaptation***. That is, should changes of the GIDS interface properties be initiated by the driver or by the system? In case of system-initiated adaptation it would also be an issue what variables can and should be utilized for decision-making. The second issue concerns ***the kind of adaptation***. Thus, which interface characteristics should be changed, and how, in order to alleviate driver workload in a GIDS system? Next, separate sections are devoted to these issues.

3.1 Human versus system initiated interface adaptation

Adaptation of the interface can either be initiated by the human operator or automatically by the system. Rouse, Geddes, and Curry (1987) emphasized the importance of ***human-initiated adaptation***. The major advantage of human-initiated adaptation is that the operator knows the status of the system. Greenstein, Arnaut, and Revesman (1986) labelled this explicit communication. Human-initiated adaptation seems especially useful for long-term adaptation since the driver can adapt the system to his or her level of experience when workload is low. Yet, for short-term adaptation human-initiated adaptation has the major disadvantage that the act of adjusting the interface to the momentary situation imposes additional workload upon the driver (Nowakowski, 1987). Also, the driver might not always foresee in time a rapid increase in workload, so adaptations may be too late.

An interesting form of human-initiated adaptation has been put forward by McKinlay (1985). He proposed to adapt the interface implicitly, that is by the possibility of the human operator to use high-level and low-level commands. In a GIDS system an extreme high-level command might be "DRIVE ME TO AMSTERDAM". The system would then carry out all necessary actions automatically. The low level extreme might involve instructing the separate controls and, thus, involves egular driving. So, by implicit adaptation, the driver can adapt the interface to his current mental workload by choosing the appropriate level of control. This prevents the need for explicit communication.

Among others, Rouse, Geddes, and Curry (1987) asserted that in some situations the human operator cannot simultaneously be coordinator and performer so ***system-initiated adaptation*** is required. Since the human operator should always know about adaptations in the system to prevent unexpected surprises explicit information about the adaptation is required. Only, with practice the human operator might learn what state the interface is in which makes explicit information unnecessary. Greenstein, Arnaut, and Revesman (1986) called this implicit communication. For example, a driver may learn that navigation information is not presented when he is engaged in critical traffic situations. Yet, the major disadvantage of system-initiated adaptation is that it reduces the amount of driver control and therewith the acceptance of the system.

Since human and system initiated adaptations are merely end points on a continuum, Endsley (1987) proposed a five step continuum:

1 the human operator initiates the adaptation (i.e. human-initiated adaptation)
2 the system proposes several alternative adaptations to the operator without executing one
3 the system proposes a specific adaptation and waits for approval
4 the system proposes an adaptation and executes it when not vetoed within a certain period of time
5 the system executes the adaptation without waiting for approval (i.e. system-initiated adaptation).

Step one and five of this continuum have already been discussed. Step two seems advisable only when driver workload is very low and can, for example, be used when the driver is setting-up the interface to his particular wishes. The third step seems also applicable in GIDS systems for it does not force the driver to respond and it keeps the driver in control. The adaptation proposal should, however, not be presented during periods of high workload to prevent driver distraction. Finally, step four may involve the risk that the driver is out of control when driver workload is high. Yet, for some minor applications it may be a nice possibility (e.g. automatic wiper control).

Since system-initiated adaptation to momentary task demands can be useful when driver workload is extremely high, a sophisticated GIDS system should be equipped with some version of it. In that case, the interface should have access to measures or indicators of driver mental workload. There are several possibilities to decide about interface adaptation; model-based decisions, decisions based on on-line measured workload indicators, and decisions based on a conjunction of both.

First, decisions can be based on a ***driver model***. The model should involve static driver variables (e.g. age, experience) but also driver mental workload in certain traffic situations under specific modifying conditions (e.g. weather conditions) and effects of prior workload (Cumming and Croft, 1973). The most simple driver model would be static, that is not changing in time, and not for a specific driver or class of drivers. In more sophisticated versions the driver model may be specific for a group of drivers (e.g. level of experience, prior history) or for a particular driver (Nowakowski, 1987). It might even be able to adapt automatically to changes of the specific driver (i.e. system-initiated long-term adaptation). The advantage of model-based adaptation is that it circumvents the problem of directly measuring driver workload. However, on-line prediction requires detailed models and fast on-board computers.

Second, decisions about interface adaptation can be based on ***on-line mental workload assessment*** (e.g. Hancock and Chignell, 1986, 1988). In the first place, one-dimensional workload measures, as discussed in section 3, may be utilized[4]. For example, Rouse, Geddes, and Curry (1987) used errors in operator performance as workload indicator and argued that an error monitor might not be able to know if a particular procedure is appropriate but that it can determine if the procedure is being executed correctly. Next, sophisticated assessment of mental workload may be multi-dimensional, that is incorporating separate measures for load on visual and auditory perception, central decision making and manual movements. The interface would then also be able to adjust workload on these dimensions separately. Problematic though, is that the only workload measure that seems applicable in future GIDS systems is some measure of driving performance since workload assessment methods that have been applied in experimental contexts are usually not feasible in a regular driving context. Although research on assessing underload from driving performance is in progress (e.g

[4] In case mental load measures consist of subjective _ timates, system-initiated adaptation is similar to human-initiated adaptation.

the DRIVE/DREAM project) the possibility of deriving driver overload from driving performance, as required for the adaptive interface, does not seem feasible.

Third and final, decisions can be based on ***mixed information***, that is information derived from both on-line measured workload and model-based workload predictions. On-line measured workload might then even be used to adapt the driver model to a specific driver. This is clearly too far-fetched for the initial GIDS systems but may be promising for the not-so-near-future.

Due to difficulties in on-line workload assessment in a regular driving situation, the predominant notion is that automatic adaptation to workload should be based on a model of the human operator and not on on-line workload measures (e.g. Rouse, Geddes, and Curry, 1987). An already sophisticated approach to model the human driver might for example start from the assumption that drivers are able to switch frequently between a strategy of closed loop (visual feedback available) and open loop (no visual feedback available) steering control (Godthelp, 1984). Another assumption might be that driving incorporates time-sharing lateral and longitudinal vehicle control (Blaauw, 1984). Wickens (1984a) and Schneider and Detweiler (1986) have mentioned a number of weaknesses of operator models in general and Rouse (1981) established that human behaviour in more complex tasks can only be modelled with an increasing loss of reliability. Recent developments in the area of cognitive modelling (Anderson, 1983; Laird, Newell, and Rosenbloom, 1987; Schneider and Detweiler, 1988 - see section ?.4) may prove to be valuable for workload prediction because these models are capable to incorporate effects of practice and time-sharing (for an initial attempt see Aasman, 1988). Distinguishing only two or three levels of driver workload by taking just a few major variables into account (e.g. kind of road, traffic density and driver experience) may already give an adequately adapting system. Alternatively, a more sophisticated model would be able to give a more-dimensional indication of workload. For example, load on visual, auditory, central-verbal, central-spatial, manual and foot 'resources'. Then, the interface should also have an indication of the load of each possible message and each driver task on each of the resources. The interface should, for instance, be capable of distinguishing between driving on a curved road (loading visual and manual resources) and negotiating a dangerous intersection (central resources are required as well).

Finally, it should be noted that model-based interface adaptation is only possible when GIDS systems are capable of recognizing momentary driving situations. The feasibility of such situation recognition will not be discussed here but poses a major challenge for any group intending to build a GIDS system.

3.2 Driver mental workload reduction by interface adaptation

Given a satisfactory method of workload prediction or assessment, the next step in interface adaptation is deciding which interface characteristics should be changed in order to alleviate operator workload. Rouse, Geddes, and Curry (1987) distinguished three methods to support the human operator: making a task easier, performing part of the task, and completely performing a task. They termed these three methods transformation, partitioning, and allocation.

Rouse, Geddes, and Curry (1987) mentioned two categories of ***transformation***. First, message or request priorities are determined and expressed in terms of the criticality of the information in question relative to goal achievement. Timeliness is supposed to be a major determinant for criticality. Once priorities are assigned, messages can be queued and presented accordingly. Second, modality and format of message presentation or of required actions can be changed. For example, a specific message may contain additional details when workload is low but exclude the details when workload has increased.

Partitioning and allocation imply that the system takes over part of a task or the whole task. The difference between the two depends merely on the definition of what constitutes a task. Since task allocation has been mentioned as a major way to reduce operator workload (e.g. Greenstein, Arnaut, and Revesman, 1986; Hancock and Chignell, 1988; Rouse, Geddes, and Curry, 1987) it will be discussed more extensively. Allocation of tasks can be static or dynamic. In ***static*** task allocation the system designer decides who is going to perform the task early in the design process. Already in 1951, Fitts attempted to characterize qualitatively those tasks performed better by machines than by humans, and those performed better by humans than machines. This list has become widely known as 'Fitts list'. Fitts list had little impact on engineering design because the criteria were overly general, nonquantitative, and incompatible (Price, 1985). Also, the list assumed that tasks would be performed by eith humans or machines alone. In a review on task allocation Price (1985) gave 11 general rules and he stated that task allocation should be embedded in the design. Price (1985) recognized the possibility that certain tasks can be performed by both humans and machines. Static task allocation should then incorporate the relative goodness of the human and that of the to-be-available machine technology. Only when human or machine performance is within acceptable limits tasks can be allocated to either.

Dynamic allocation of tasks or tasks has been pursued among others by Rouse and Chu (e.g. Chu and Rouse, 1979; Rouse, 1981). Dynamic allocation entails that sometimes the operator and sometimes the machine performs a particular task. Rouse (1981) stated that a dynamic approach to task allocation has numerous advantages including better use of the system's resources, less variability of the operator's workload and the possibility for the operator to have an improved knowledge of the overall system state. The basic idea is that a particular task is allocated to the controller (human or computer) that has at that moment the most resources available for performing the task. According to Rouse (1981) a major problem is the ambiguity of who is in command. This may even occur when the adaptation has been initiated by the human operator since he may forget how he has set up the system. Also, because the interface does not behave constant over time (see Chapter 2) effects of practice will be reduced. Confusion may even lead to a situation where the operator and the machine try to adapt to one another (Edmonds, 1981). So, in cases when mental workload is moderately so that reallocation is not necessary, reallocation should not be used at any time (Hopkin, 1975). Finally, the process of task reallocation should not be disruptive (Chignell and Hancock, 1985) for a large number of sudden, discrete changes might lead to a worsening, rather than an improvement in performance (see section 3.3 on this issue).

Lastly, adaptation to driver experience (i.e. long-term adaptation) should be handled cautiously. Care should be taken that working under one interface mode of experience yields positive transfer to working in a more experienced mode. Possibly, in a novice mode of operation simple actions are learned that are also present in the expert mode of operation. When these actions can be applied fast and accurately the driver should switch to a higher interface level of experience which assumes that such actions can be carried out in rapid succession or in more attention demanding situations.

3.3 The electric cocoon

Following a number of dramatic accidents and incidents as a result of flight-deck automation a more critical attitude has developed toward automation of tasks (Wiener and Curry, 1980; Wiener, 1985). Instead of progressing automation Wiener and Curry (1980) proposed the concept of an "electric cocoon". This would involve a system that allows pilots all freedom they want and thus it would always involve a human-initiated system. Only in case certain limits are crossed, pilots find themselves surrounded by a multidimensional warning and alerting system that informs the crew when they are crossing some limit and what to do. In addition, Wiener and Curry (1980) recommended trend alarms to prevent warnings when limits have already been reached. They proposed that future research on flight-deck automation should be aimed at bringing the crew back into the loop and prevent them to be merely monitors. Automation might better be used to allow the crew to create their own interface by setting critical limits and other system characteristics. Also, the computer can better monitor the pilot than vice versa.

The lesson for the design of GIDS systems is that system-initiated actions should be applied with extreme care and that drivers should, preferably, be allowed to create their own interface. Driving should not become a bore because almost everything is taken care of automatically and the driver only monitors the system, nor should it be an annoyance since the driver is out of control. Rather the GIDS system should give drivers a feeling of freedom as long as they stay within the safety limits (Rouse, 1988). So, making a system that motivates the driver to work with while at the same time increasing traffic safety is a major challenge for any group that attempts to design a GIDS system.

3.4 Recommendations

The review on adaptive interfaces allows the following recommendations for GIDS systems to be made:

1 System-initiated interface adaptation should only be used when immediate danger requires all attention of the driver (i.e in short-term adaptation). In all other cases driver-initiated adaptation of the interface is preferred although the system may propose adaptations that require approval of the driver. As an alternative

to system-initiated short-term interface adaptations, the driver may also be protected by a sophisticated warning and alert system, including trend alarms. Warnings should be self-explaining in order to minimize workload and should prevent fright reactions. An elegant way of human-initiated interface adaptation is implicit adaptation. This implies that the driver is capable of controlling the car on several hierarchical levels which reduces the need for explicit communication.

2 Long-term interface adaptation should preferably also be human-initiated and should allow drivers to create an interface adapted to their own wishes. The system may also make proposals when the driver is changing the interface. These proposals should be based on research that has indicated how transfer of training from the novice mode to the expert driver mode is maximal.

3 No matter who initiates the adaptation of the interface, any adaptation should be clearly indicated (e.g. by the colour of the instrument lights). Also, usual driver actions should be allowed to be executed in the normal way - only the interface presentation format may change, not the way it is controlled.

4 Interface adaptation should involve postponing messages that load the driver. Also, information can be presented in another modality or in a changed format. However, care should be taken with such adaptations - even when human-initiated - since unexpected changes or messages in an unfamiliar format, for example, in a mentally loading situation, may increase driver workload even more. Most probable candidates of driver tasks to be changed are those tasks that do not improve much with experience. When driver tasks are changed that are highly practice, workload will increase considerably.

5 The successful introduction of adaptive interfaces will inevitably depend on user understanding and acceptance. This implies that ambiguity and lack of operator control should be prevented in any case. On the other hand, the user may become over-reliant and will not even consider the possibility of a failure. GIDS systems should be designed to prevent such misunderstandings.

6 For any intelligent GIDS system, whether equipped with a sophisticated warning system or a system-initiated interface, a model of the driver is required which is capable of predicting driver

workload as a function of environmental, traffic, and driver variables. Research on this issue is badly required. Also, research is required to determine the load of each task or message in a GIDS system, preferably on separate dimensions of workload.

4 EPILOGUE

4.1 Summary

This report examined issues concerning mental load which are relevant for GIDS-like intelligent interfaces. It was concluded that humans are not able to perform several tasks at the same time when they are unpractised or when one of the tasks demands all attention. In such cases humans execute the tasks by shifting attention between them. In other multitask situations, especially when one of the tasks is of a continuous nature, multiple resource theory and the S-C-R principle make valuable predictions regarding optimal timesharing of tasks. It has been emphasized that more consistent tasks allow for greater effects of practice and alleviate workload. Under certain conditions (e.g. consistent order and timing) separate tasks may even integrate and the situation ceases to be multitask. Otherwise, a mental model of the task environment enables the human operator to expect and prepare specific actions.

Some general problems with the concept of workload were outlined. It was established that workload can be predicted by means of an operator model or assessed by a number of methods. Four general methods of workload assessment were introduced, primary task performance, secondary task performance, physiological measures and subjective estimates of workload. It was concluded that for general workload evaluation in an experimental driving situation a combination of subjective estimation and the 0.1 Hz. cardiovascular component are well suited. For load on specific resources, the secondary task technique or resource-specific physiological measures are also appropriate. Yet, in secondary task situations care should be taken to prevent deterioration of the primary task (i.e. driving).

Next, literature on adaptive interfaces has been considered. Aim of such interfaces is to alleviate operator workload by adapting the car interface when workload is high, for instance by ceasing low priority tasks and complex or unnecessary messages. This is short-term adaptation. Long-term adaptation implies that the interface adapts to the operator level of experience. Two major issues have been discussed. The first concerned who takes the initiative to adapt the interface, the operator or the machine. The second issue was about the ways to alleviate driver workload.

In order to prevent performance decrements, e.g. by ambiguity about who is in control, it seemed that only under extreme conditions the interface may change its properties automatically. Another reason for reducing application of system-initiated adaptation was the increased need for communication although this may become obsolete when the operator gets to know the system well. An interesting possibility was to give the operator the opportunity to use both low- and high-level commands. Next, there was the issue of how the GIDS system can detect driver overload. Regular methods of assessing workload do not seem to be applicable in regular cars. More or less cognitive driver models for workload estimation may be a more feasible and practical solution for assessing driver workload in a GIDS system.

The second issue in adaptive interfacing is how to change the interface when workload is too high. Three groups of possible changes were proposed, transformation, partitioning, and allocation. Transformation involves simplifying the task by, for example, reducing the information content or by postponing some tasks. Partitioning and allocation imply that some tasks are taken over by the machine. Yet, the latter is only possible when both the human operator and the machine are capable of executing these tasks satisfactory. System-initiated adaptation should never be executed at levels of workload that are not excessive. This would lead to erosion of skills, boredom, complacency, reduced failure-detection skills and reduced skill acquisition. Lastly, care should be taken that driving in a novice driver interface mode yields positive transfer of training to driving in an experienced driver mode.

4.2 Issues for future research

Three major areas for further research have been established. First, driver workload should be assessed quantitatively as a function of various driver, traffic and environmental variables. These variables should be incorporated in the driver model that is used for system-initiated adaptation. Also, the workload of all messages to the driver and all possible driver tasks should be assessed.

A second area where knowledge is lacking is the area of driver task design. That is, how should the new driver tasks look like in order to minimize mental workload when performed concurrently with other driving tasks. Variables of interest seem to be consistency of order and timing, and stimulus and response modalities.

Finally, research should address the transfer of training issue. In other words, what characteristics of the interface should remain unchanged when switching from a novice mode to an expert mode of operating and still have maximal transfer?

REFERENCES

Aasman, J. (1988). Implementations of car-driver behaviour and psychological risk models. In T. Rothengatter, & R. de Bruin (Eds.), Road user behavior: Theory and research (pp. 106-118). Assen: van Gorcum.

Aasman, J., Mulder, G., & Mulder, L.J.M. (1987). Operator effort and the measurement of heart-rate variability. Human Factors, 29, 161-170.

Adams, J.A. (1984). Learning of movement sequences. Psychological Bulletin, 96, 3-28.

Allport, D.A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. Quarterly Journal of Experimental Psychology, 24, 225-235.

Anderson, J.R. (1983). The architecture of cognition. London: Harvard University Press.

Bauer, L.O., Goldstein, R., & Stern, J.A. (1987). Effects of information- processing demands on physiological response patterns. Human Factors, 29, 213-234.

Biesta, P.W., & Blaauw, G.J. (1976). The effect of environmental factors on measures for workload and driving behaviour (in Dutch). Report IZF 1976-7. Soesterberg, The Netherlands: TNO Institute for Perception.

Blaauw, G.J. (1984). Car driving as a supervisory control task. Unpublished doctoral dissertation. Soesterberg, The Netherlands: TNO Institute for Perception.

Broadbent, D.E. (1958). Perception and Communications. London: Pergamon Press.

Broadbent, D.E. (1982). Task combination and selective intake of information. Acta Psychologica, 50, 253-290.

Brown, T.L., & Carr, T.H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance, 15, 686-700.

Carswell, C.M., & Wickens, C.D. (1987). Information integration and the object display. An interaction of task demands and display superiority. Ergonomics, 30, 511-527.

Chignell, M.H., & Hancock, P.A. (1985). Knowledge-based load leveling and task allocation in human-machine systems. Proceedings of the Annual Conference on Manual Control, 21, 9.1-9.11.

Chu, Y.Y., & Rouse, W.B. (1979). Adaptive allocation of decision making responsibility between human and computer in multi-task situations. IEEE Transactions on Systems, Man, and Cybernetics, SMC-9, 12, 769-778.

Colle, H.A., Amell, J.R., Ewry, M.E., & Jenkins, M.L. (1988). Capacity equivalence curves: A double trade-off curve method for equating task performance. Human Factors, 30, 645-656.

Commission of the European Communities (1989). DRIVE '89. Brussels: Commission of the European Communities, DG XIII, April 1989.

Cumming, R.W., & Croft, P.G. (1973). Human information processing under varying task demand. Ergonomics, 16, 581-586.

Curry, G.A., Hieatt, D.J., & Wilde, G.J.S. (1975). Task load in the motor vehicle operator: A comparative study of assessment procedures. Ottawa: Ministry of Transport, CR7504.

Deutsch, J.A., & Deutsch, D. (1963). Attention: Some theoretical considerations. Psychological Review, 70, 80-90.

Edmonds, E.A. (1981). Adaptive man-computer interfaces. In M.J. Coombs, & J.L. Alty (Eds.), Computing Skills and the User Interface (pp. 389-426). London: Academic Press.

Eggemeier, F.T., & Stadler, M.A. (1984). Subjective workload assessment in a spatial memory task. Proceedings of the Human Factors Society, 29, 680-684.

Endsley, M.R. (1987). The application of human factors to the development of expert systems for advanced cockpits. Proceedings of the Human Factors Society, 31, 1388-1393.

Fisk, A.D., Ackerman, P.L., & Schneider, W. (1987). Automatic and controlled processing theory and its applications to human factor problems. In P.A. Hancock (Ed.), Human Factors Psychology (pp. 159-197). Amsterdam: North Holland.

Fitts, P.M. (Ed.).(1951). Human engineering for an effective air-navigation and traffic-control system. Columbus, OH: Ohio State University Research Foundation.

Galer, M. & Simmonds, G. (1984). Ergonomic aspects of electronic instrumentation: A guide for designers SP-576. Warrendale, PA: Society of Automotive Engineers, Inc.

Gladstones, W.H., Regan, M.A., & Lee, R.B. (1989). Division of attention: The single-channel hypothesis revisited. Quarterly Journal of Experimental Psychology, 41A, 1-17.

Godthelp, H. (1984). Studies on human vehicle control. Unpublished doctoral dissertation. Soesterberg, The Netherlands: TNO Institute for Perception.

Greenstein, J.S., Arnaut, L.Y., & Revesman, M.E. (1986). An empirical comparison of model-based and explicit communication for dynamic human- computer task allocation. International Journal for Man-Machine Studies, 24, 355-363.

Gstalter, H. (1985). Workload and stress in driving. In I.D. Brown, R. Goldsmith, K. Coombes, & M.A. Sinclair (Eds.), Proceedings of the ninth congress of the International Ergonomics Association (pp. 994-996). London: Taylor & Francis.

Hancock, P.A. (1988). The effect of gender and time of day upon the subjective estimate of mental workload during the performance of a simple task. In P.A. Hancock, & N. Meshkati (Eds.). Human mental workload. Amsterdam: North-Holland, 239-250.

Hancock, P.A. (1989). The effect of performance failure and task demand on the perception of mental workload. Applied Ergonomics, 20, 197-205.

Hancock, P.A., & Chignell, M.H. (1986). Toward a theory of mental workload: Stress and adaptability in human-machine systems. Proceedings of the Conference on System, Man, and Cybernetics, 378-383.

Hancock, P.A., & Chignell, M.H. (1988). Mental workload dynamics in adaptive interface design. IEEE Transaction on Systems, Man, and Cybernetics, 18, 647-658.

Hancock, P.A., Chignell, M.H., & Loewenthal, A. (1985). An adaptive human-machine system. Proceedings of the IEEE Conference on System, Man, and Cybernetics, 627-630.

Hancock, P.A., Rodenberg, G.J., Mathews, W., & Vercruyssen, M. (1988). Estimation of duration and mental workload at differing times of the day by males and females. Proceedings of the Human Factors Society, 32, 857-861.

Hart, S.G. (1986). The relationship between workload and training: An introduction. Proceedings of the Human Factors Society, 30, 1116-1120.

Hart, S.G., & Bortolussi, M.R. (1984). Pilot errors as a source of workload. Human Factors, 26, 545-556.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock &, N. Meshkati (Eds.), Human Mental Workload (pp. 139-183). Amsterdam: North-Holland.

Hicks, T.G., & Wierwille, W.W. (1979). Comparison of five mental workload assessment procedures in a moving-base simulator. Human Factors, 21, 129-143.

Hirst, W., Spelke, E.S., Reaves, C.C., Caharak, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. Journal of Experimental Psychology: General, 109, 98-117.

Hopkin, V.D. (1975). The controller versus automation. In A. Benoit (Ed.), A survey of modern air traffic control, AG-209 (pp. 43-59). Neuilly sur Seine, France: AGARD.

Jahns, D.W. (1973). A concept of operator workload in manual vehicle operations. Forschungsbericht 14. Meckenheim, FRG: FAT.

Janssen, W.H. (1979). Routeplanning and guidance: a literature study (in Dutch). IZF Report 1979 C-13. Soesterberg, The Netherlands: TNO Institute for Perception.

Janssen, W.H., & Gaillard, A.W.K. (1984). Task load and stress on the road: Preliminaries to a model of route choice. IZF Report 1984 C-10. Soesterberg, The Netherlands: TNO Institute for Perception.

Johanssen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., & Wickens, C. (1979). Final report of experimental psychology group. In N. Moray (Ed.), Mental Workload. Its Theory and Measurement (pp. 101-114). New York: Plenum Press.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, NJ: Prentice-Hall.

Kramer, A.F., Sirevaag, E.J., & Braune, R. (1987). A psychophysiological assessment of operator workload during simulated flight mission. Human Factors, 29, 145-160.

Laird, J., Newell, A., & Rosenbloom, P. (1987). Soar: An Architecture for General Intelligence. AI, 33, 1-64.

Logan, G.D. (1985). Skill and automaticity: Relations, implications, and future directions. Canadian Journal of Psychology, 39, 367-386.

MacDonald, W.A. (1979). The measurement of driving task demand. Melbourne: Corvina.

McCormick, E.J., & Sanders, M.S. (1982). Human Factors in Engineering and Design. Auckland: McGraw-Hill.

McKinlay, W.H. (1985). Automation and pilot interactions in night or all-weather technical operations. In Guidance-Control-Navigation Automation for Night All-Weather Tactical Operations, CP-387 (pp. 8.1-8.10). Neuilly sur Seine, France: AGARD.

McLeod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. Quarterly Journal of Experimental Psychology, 29, 651-667.

Miller, G.A. (1956). The magic number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Michon, J.A. (1966). Tapping regularity as a measure of perceptual motor load. Ergonomics, 9, 401-412.

Moray, N. (1967). Where is capacity limited? A survey and a model. Acta Psychologica, 27, 84-92.

Moray, N. (1979). Models and measures of mental workload. In N. Moray (Ed.), Mental Workload. Its Theory and Measurement (pp. 13-22). New York: Plenum Press.

Moray, N., Turksen, B., Aidie, P., Drascic, D., Eisen, P., Kruschelnicky, E., Money, L., Schonert, H., & Thornton, C. (1986). Progress in mental workload measurement. Proceedings of the Human Factors Society, 30, 1121-1123.

Mulder, G. (1973). Mental load and the measurement of heart rate variability. Ergonomics, 16, 69-83.

Mulder, G. (1979). Sinusarrythmia and mental workload. In N. Moray (Ed.), Mental workload: Its theory and measurement (pp. 327-344). New York: Pergamon Press.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer, & A.F. Sanders (Eds.), Perspectives on perception and action (pp. 361-394). Hillsdale, NJ: Erlbaum.

Nowakowski, M. (1987). Personalized recognition of mental workload in human-computer systems. In G. Salvendy, S.L. Sauter, & J.J. Hurrell (Eds.), Social, ergonomic and stress aspects of work with computers (pp. 75-82). Amsterdam: Elsevier.

Noy, Y.I. (1987). Theoretical review of the secondary task methodology for evaluating intelligent automobile displays. Proceedings of the Human Factors Society, 31, 205-209.

Price, H.E. (1985). The allocation of functions in systems. Human Factors, 27, 33-45.

Reid, G.B. (1985). Current status of the development of the Subjective Workload Assessment Technique. Proceedings of the Human Factors Society, 29, 220-223.

Roscoe, A.H. (1987). Introduction. In A.H. Roscoe (Ed.), The practical assessment of pilot workload, AG-282 (pp. 1-10). Neuilly sur Seine, France: AGARD.

Rouse, W.B. (1981). Human-Computer interaction in the control of dynamic systems. Computing Surveys, 13, 71-99.

Rouse, W.B. (1988). Adaptive aiding for human/computer control. Human Factors, 30, 431-443.

Rouse, W.B., Geddes, N.D., & Curry, R.E. (1987). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. In T.P. Moran (Ed.), Human-Computer Interaction (pp. 87-122). Hillsdale, NJ: Erlbaum.

Sanders, A.F. (1979). Some remarks on mental load. In N. Moray (Ed.), Mental Workload. Its Theory and Measurement (pp. 41-78). New York: Plenum Press.

Sanders, A.F. (1983). Towards a model of stress and human performance. Acta Psychologica, 53, 61-97.

Sayers, B. (1973). Analysis of heart rate variability. Ergonomics, 16, 17-33.

Schneider, W., & Detweiler, M. (1986). Changes in performance in workload with training. Proceedings of the Human Factors Society, 30, 1128-1133.

Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/-control architecture. Human Factors, 30, 539-566.

Schneider, W., & Fisk, A.D. (1984). Automatic category search and its transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 1-15.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Shaffer, L.H. (1975). Multiple attention in continuous verbal tasks. In P.M.A. Rabbitt, & S. Dornic (Eds.), Attention and Performance V (pp. 157-167). New York: Academic Press.

Shladover, S.E. (1988). Roadway automation technology - research needs. Transportation Research Board, 68th Annual Meeting, January 22-26, Washington, D.C., paper no. 880208.

Smiley, A., & Michon, J.A. (1989). Conceptual Framework for Generic Intelligent Driver Support. Deliverable No. 1 GIDS-GEN. Groningen: Traffic Research Centre, University of Groningen.

Taylor, D.H. (1964). Drivers' galvanic skin response and the risk of accident. Ergonomics, 7, 439-451.

Ursin, H., & Ursin, R. (1979). Physiological indicators of mental workload. In N. Moray (Ed.), Mental Workload. Its Theory and Measurement (pp.349-366). New York: Plenum Press.

Verwey, W.B. (1989). Simple in-car route guidance from another perspective: Modality versus coding. Proceedings of the first Vehicle Navigation and Information Systems Conference (VNIS'89), Toronto 1989, (pp. 56-60). Toronto: IEEE.

Verwey, W.B. (In preparation). Integration of sequential actions and the role of preparation. TNO report in preparation.

Vicente, K.J., Thornton, D.C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: a measure of mental effort. Human Factors, 29, 171-182.

Welford, A.T. (1988). Acquiring skill: Some experiences. Journal of Motor Behaviour, 20, 458-461.

Wetherell, A. (1981). The efficacy of some auditory-vocal subsidiary tasks as measure of the mental load on male and female drivers. Ergonomics, 24, 197-214.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson (Ed.). Attention and Performance VIII (pp. 239-257). Hillsdale, NJ: Erlbaum.

Wickens, C.D. (1984a). Engineering Psychology and Human Performance. Columbus: Merill.

Wickens, C.D. (1984b). Processing resources in attention. In R. Parasuraman, & D.R. Davies (Eds.), Varieties of attention (pp. 63-102). London: Academic Press.

Wickens, C.D., Fracker, L., & Webb, J. (1987). Cross-modal interference and task integration: Resources or preemption switching? Proceedings of the Human Factors Society, 31, 679-682.

Wickens, C.D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification, Human Factors, 30, 599-616.

Wickens, C.D., Sandry, D., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, output, and central processing. Human Factors, 25, 227-248.

Wiener, E.L. (1985). Beyond the sterile cockpit. Human Factors, 27, 75-90.

Wiener, E.L., & Curry, R.E. (1980). Flight-deck automation: Promises and problems. Ergonomics, 23, 955-1011.

Wierwille, W.W., & Gutmann, J.C. (1978). Comparison of primary and secondary task measures as a function of simulated vehicle dynamics and driving conditions. Human Factors, 20, 233-244.

Williges, R.C., & Wierwille, W.W. (1979). Behavioral measures of aircrew mental workload. Human Factors, 21, 549-574.

Willingham, D.B., Nissen, M.J., & Bullemer, P. (1989). On the development of procedural knowledge. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 1047-1060.

Yeh, Y.Y., & Wickens, C.D. (1988). Dissociation of performance and subjective measures of workload. Human Factors, 30, 111-120.

Soesterberg, January 26, 1990

Drs.Ing. W.B. Verwey

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) TD 90-0018	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER IZF 1990 B-3
4. PROJECT/TASK/WORK UNIT NO. 735.4	5. CONTRACT NUMBER B89-59	6. REPORT DATE January 26, 1990
7. NUMBER OF PAGES 40	8. NUMBER OF REFERENCES 98	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

Adaptable driver-car interfacing and mental workload: A review of the literature

11. AUTHOR(S)

W.B. Verwey

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Institute for Perception
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

TNO Division of National Defence Research
Koningin Marialaan 21
2595 GA DEN HAAG

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

This report discusses a number of issues that are important for the design of an adaptive interface or dialogue controller in a car equipped with a GIDS system. There are two major sections, each followed by recommendations for interface design.
In the first section (Chapter 2), a short review of human performance in multitask situations is given. Practice is assumed to be of major significance for mental workload reduction especially in tasks that are executed concurrently with similar tasks. Also, issues concerning the concept, prediction and assessment of mental workload are discussed. A distinction is made between mental workload assessment for the evaluation of GIDS systems and for the adaptation of the interface to the driver workload. Recommendations are given for workload assessment in evaluation studies.
The second section of the report (Chapter 3) discusses literature on adaptive interfaces. Distinctions are made between human and system-initiated interface adaptation and between short-term and long-term adaptation. The section ends with recommendations regarding when and how to adapt the interface characteristics to driver workload.
Finally, the report concludes with a summary and issues for further research in chapter 4.

16. DESCRIPTORS	IDENTIFIERS
Driving Workload Displays	Adaptable Interfacing

17a. SECURITY CLASSIFICATION (OF REPORT) -	17b. SECURITY CLASSIFICATION (OF PAGE) -	17c. SECURITY CLASSIFICATION (OF ABSTRACT) -
18. DISTRIBUTION/AVAILABILITY STATEMENT Unlimited availability		17d. SECURITY CLASSIFICATION (OF TITLES) -

VERZENDLIJST

1. Hoofddirecteur van de Hoofdgroep Defensieonderzoek TNO

2. Directie Wetenschappelijk Onderzoek en Ontwikkeling Defensie

3. { Hoofd Wetenschappelijk Onderzoek KL
Plv. Hoofd Wetenschappelijk Onderzoek KL

4,5. Hoofd Wetenschappelijk Onderzoek KLu

6. { Hoofd Wetenschappelijk Onderzoek KM
Plv. Hoofd Wetenschappelijk Onderzoek KM

7. Wnd. Hoofd Afd. Militair Geneeskundig Beleid

8. Inspecteur Geneeskundige Dienst KL
Brig.Gen.-arts B.C. Mels

9. Inspecteur Geneeskundige Dienst KLu
Cdre J.Th. Versteeg

10. Inspecteur Geneeskundige Dienst Zeemacht
Cdr-arts A.J. Noordhoek

11, 12, 13. Hoofd van het Wetensch. en Techn. Doc.- en Inform. Centrum voor de Krijgsmacht

LEDEN WAARNEMINGS CONTACT COMMISSIE

14. Maj.Ir. W.C.M. Bouwmans
15. LTZAR1 F.D.J.R. Feunekes
16. Dr. N. Guns
17. Drs. C.W. Lamberts
18. Ir. P.H. van Overbeek
19. Drs. W. Pelt
20. Maj. dierenarts H.W. Poen
21. Drs. F.H.J.I. Rameckers
22. LTZSD2OC KV Drs. M.B.A.M. Scheffers
23. Prof.Ir. C. van Schooneveld
24. LKol.Drs. H.W. de Swart
25. Ir. M. Vertregt
26. Kol. vliegerarts B. Voorsluijs

Extra exemplaren van dit rapport kunnen worden aangevraagd door tussenkomst van de HWOs of de DWOO.
