Over the past thirty years, the difficult tasks that operators, in particular aircraft pilots and air traffic control operators, have had to perform have drawn attention to the area of mental workload. General questions have been asked such as "How busy is the operator?", "How many tasks can he handle safely?" and "Does the operator have to ‘try hard’ to maintain an adequate level of performance?". If task demands are high in relation to the operator’s capabilities, errors may occur, and in interaction with neglected classical human factors issues such as a proper layout of instrumentation panels, these errors may become critical for safety. Even economic interest can raise workload-related questions. As an example, Wickens (1992) described a controversy between an airline industry and a pilot association. The airline industry claimed that a certain class of aeroplanes could be flown by two crew members, while the pilot organization claimed that demands at peak times would be excessive and would require a three person complement. Such issues have called for a definition of mental workload and the methods to assess it.

This thesis is about how to measure driver workload. In the present chapter the more theoretical aspects of mental workload in general and driver mental workload in particular will be introduced. In chapter 2 a model that relates task demands to workload and performance will be presented. In chapter 3 and 4, the general criteria for workload measurement techniques are described, followed by a categorization of measures. Properties of different measurement techniques and experience from non-traffic research will be reported in chapter 4. From chapter 5 onwards the focus is on the use of the techniques in traffic research. Although some of the techniques have been applied in traffic research, an overview and review of their characteristics in this specific field is missing. Driving is a very dynamic task in a changing environment. Moreover, contrary to many laboratory tasks, the driving task is to a large extent influenced by drivers themselves. The driver’s influence on the task ranges from strategic aspects such as route selection, to ‘control behaviour’ such as the accuracy in lane keeping. In particular, the increase in RTI (Road Transport Informatics) makes the evaluation of mental workload techniques for use in traffic research relevant and urgent. With an increase in in-vehicle RTI applications, road safety may be negatively affected. Much is to be gained by thorough evaluation of the mental load effects of new equipment before introduction to the market (e.g., Parkes, 1991). In chapter 5, the measurement techniques will be evaluated on sensitivity, reliability and operational aspects on the basis of results of several field studies. Sensitivity and dissociation of different measures will also be evaluated in the context of the mental workload model presented in chapter 2. The so-called ‘workload
redline’, which indicates the critical level of too much mental workload, will be linked to the model and its potential, as well as the problems associated with correct redline determination, will be discussed.

Theories relevant for mental workload

In order to understand mental workload, the introduction of some basic concepts is required. The concept of a limited processing capacity can be found in many theories (e.g., Broadbent 1958, Kahneman, 1973, Posner, 1978, Wickens, 1984). Kahneman (1973) specifies the metaphor of a single undifferentiated capacity (the ‘modal’ view) from which resources are available for task performance. O’Donnell & Eggemeier (1986) make no difference between the metaphoric words ‘capacity’ and ‘resource’ and use the words as interchangeable terms. Wickens (1992) disagrees with this. He defines capacity as the maximum or upper limit of processing capability, while resources represent the mental effort supplied to improve processing efficiency. This is in line with Norman & Bobrow (1975) who also refer to resources as processing effort. In this thesis the differentiation between capacity as upper limit of capability and resources as amount of processing facilities allocated will be followed. Resources are characterized by two general properties: their deployment is under voluntary control and they are scarce. Only the very simple resource models consider capacity to be fixed. According to Kahneman (1973) there is some elasticity in capacity and the availability of resources, the mobilization of resources could be increasingly possible, e.g. as a result of increased processing load.

The relation between resource allocation and task performance is supposed to be linear, until the moment all resources are invested. From that point on, no more resources can be invested and task performance will remain stable. Norman & Bobrow (1975) call such a task resource-limited. The resource-limited task is opposed to a data-limited task. When performing a data-limited task, additional available resource investment does not lead to increased performance due to limitations in data quality. Although the theory could be applied to a variety of situations, it could not explain why effective time-sharing and unaffected performance could occur when a second auditory task was added to a primary visual task.

In the 1980’s Wickens proposed a multiple-resource theory in which different resources for different modalities are assumed (Wickens, 1984). Most prominent are the auditory and visual resources. In addition to these, central resources are supposed, which are required for the performance of almost all tasks. An overlap in resource requirement, e.g. the performance of two auditory tasks, soon requires full auditory capacity use. In that case performance on both tasks will be affected. Tasks that require different resources, e.g., a visual task combined with an auditory task, will not directly interfere with each
other and performance of either task can remain unaffected, provided there is no performance decrement caused by central resource use.

The concept of multiple resources is connected to three dimensions. The first dimension is the processing stage, i.e. perception (including encoding), central, and response processing. The second dimension is modality of input and response. The auditory, visual and tactile modality draw upon different resources and cross-modal time-sharing can be better performed than intramodal timesharing. Listening to someone and watching something at the same time associate better than listening to two things at the same time. The third dimension is the processing code. The processing code can be either verbal or spatial.

With respect to stage, the multiple-resource theory predicts more interference between tasks if both tasks demand spatial processes, or if both demand verbal processing across any stage. So, even if the perceptual modality is different (e.g. auditory and visual) the tasks will interfere if both require (e.g.) verbal central processing. The second dimension, separateness of modality resources, was later dropped by Wickens (1991), mainly due to the influence of physical restrictions. Two competing visual channels cannot be watched at the same time and hence require scanning, an additional cost. Moreover, two simultaneously presented auditory messages will mask one another. In the last dimension, different codes can be better combined. A manual, spatial, process can, for instance, be successfully time-shared with a visual process. A well known example is typing and sight-reading.

Capacity theories have been linked to computational processes and to energetical mechanisms (G.Mulder, 1986). In the processing of information from information uptake to overt or covert reaction, a series of stages are passed in which computational processes are performed. At least four stages of processing are identified (e.g., Sanders, 1983); stimulus preprocessing, feature extraction, response choice and response adjustment. Each stage is related to a processing module with a limited capacity. A large number of these processes are not conscious. These processes are fast and automatic and cannot be subjectively assessed (Meijman & Mulder, 1992). There are, however, other processes that require working memory and are (partly) conscious and can be subjectively determined. These two classes of processes have often been labelled automatic resp. controlled processing (see below). Electrical and magnetical brain activity during the performance of information processing tasks can help to identify which brain mechanisms are mobilised in different stages of information processing (e.g., Brookhuis, 1989, Wijers, 1989).

Energetical mechanisms facilitate the availability of computational processes, and depend upon the mental or physical state of the individual. Three energetical resources have been identified (Pribram & McGuiness, 1975); arousal, activation and a compensatory resource labelled ‘effort’. Note that the resource is labelled effort; this
should not be confused with the allocation of resources that Norman & Bobrow (1975) indicated as processing effort. The effort mechanism is active in the case of attention demanding information processing, or in the case that the operator’s state differs too much from the required state. This last condition has been put forward by Hockey in his State Control Theory (Hockey, 1986). According to this theory, central executive mechanisms compare the current cognitive state with a required or target state. Whenever there is a mismatch between these two states the energetical construct of effort can be involved in actively manipulating the current state towards the target state. Hockey calls such a manipulation ‘state management’. By investing mental effort the detrimental influences of stressors (such as noise, information overload or monotony) can be successfully counteracted. A task of a highly monotonous nature, for instance, may stimulate compensatory mental effort to maintain performance. In his state-control theory, Hockey (1986) also puts forward the aspect of strategy. A minimal strategy for example is one of inaction. Performance will probably not be very high, while the effort costs are always low. Another option is that, instead of adapting the current state, different criteria for optimal performance are accepted. However, this type of goal changes often result in decreased performance. A last option is to deal directly with the source of environmental influence. A window can be opened in order to regulate environmental temperature, or it can be closed in order to reduce the noise level (Van Ouwerkerk et al., 1994a).

Sanders (1983) and G.Mulder (1986) have put forward an integration of energetic and computational models. In this model the efficiency of computational processes is affected by the energetical resources: arousal, activation and effort. Arousal affects feature extraction, while activation affects the motor organization. In tasks that require retention of information in the Working Memory, the effort mechanism is the structure that supplies energy for processing. G.Mulder (1986) assumes that there are two forms of effort: effort for tasks that require controlled information processing (computational effort), and effort in the case that an individual has to change the current energetical resource state towards a required state (compensatory effort). Cnossen (1994) labels the first as task-related effort and the latter as state-related effort.

In the information processing and task performance literature two types of theories dominate; physiological theories and cognitive theories. Quite often adherents of these two theories make use of the same terminology, which very much complicates understanding. Sometimes resources are referred to as processing modules with a limited capacity, while at other times, resources are referred to as physiological energetical structures. Nevertheless, Sanders (1983) and G.Mulder (1980, 1986) have made clear that these two types of theories are not mutually exclusive, and have proposed an integration. In the
following paragraph the concepts that are important for the measurement of mental workload will be described. Links to both cognitive and physiological theories remain apparent.

**The concept of mental workload and its assessment**

A simplistic definition of workload is that it is a demand placed upon humans. This definition attributes workload exclusively to an external source. An indication of workload, however, can be better defined in terms of experienced load. With experienced load, workload is not only task-specific, it is also person-specific (Rouse et al., 1993). Not only individual capabilities, but also motivation to perform a task, strategies applied in task performance, as well as mood and operator state, affect experienced load. In the (mental) workload literature, task demands and the effect of these demands on the operator are sometimes, unfortunately, indicated with the same term, ‘workload’. For reasons of clarity in this text demand will henceforth be used to indicate the task demands. Demand is determined by the goal that has to be attained by means of task performance, and is, once the goal has been set, external and independent of the individual. Load or workload will be used to describe the effect the demand has on the operator in terms of stages that are used in information processing and their energetics. More specifically, workload is the specification of the amount of information processing capacity that is used for task performance. In the concept of mental workload how the goal is reached (e.g. the order of actions) and individual restrictions imposed upon performance (e.g. in terms of accuracy or speed) are included. Therefore workload depends upon the individual, and owing to the interaction between operator and task structure, the same task demands do not result in an equal level of workload for all individuals. Directly related to demand is (task) complexity. Complexity increases with an increase in the number of stages of processing that are required to perform a task. Task demand and complexity are mainly external, but both depend upon (subjective) goals set for task performance. Difficulty of a task is related to the processing effort (amount of resources) that is required by the individual for task performance, and is dependent upon context, state, capacity and strategy or policy of allocation of resources. Kantowitz (1987) has proposed this differentiation between complexity and difficulty as a property of, respectively, the task in isolation versus the interaction between task and individual. The parallel with, e.g., a maths exam is noticeable; the goal that has to be reached, solving the mathematical questions, is the same for everyone and depends upon the number of calculations that have to be performed. However, goal setting affects the task demands, and there is a difference between ‘considering a C sufficient versus going for an A’. How difficult the calculations are depends very much upon the individual who has to perform the calculations. They may be relatively easy for a trained or experienced person and very hard for a novice. After a sleepless night,
however, the task will be more difficult even for the experienced person.

O’Donnell & Eggemeier (1986) define workload as that portion of the operator’s limited capacity that is actually required to perform a particular task. Workload measurement is the specification of the amount of capacity used. In this definition also, workload is not solely task-centred. Mental workload depends upon the demands in relation to the amount of resources the operator is willing or able to allocate, and is therefore a relative concept (Meijman & O’Hanlon, 1984, Zijlstra & Mulder, 1989).

In workload measurement, not only processing effort or resource allocation (Norman & Bobrow, 1975) are of primary importance, the term effort is also used for the mobilisation of additional resources as a compensatory process (see G.Mulder, 1980, Aasman et al., 1987, Vicente et al., 1987). Effort reflects the operator’s reaction to demand and the amount of effort being expended is considered by many to be one of the most important components of (if not equal to) mental workload. Vicente et al. (1987) mention two important reasons for this. Firstly, the effort expended by the operator is not necessarily related to input load (demand). The operator’s reaction to the demand depends on internal goals and adopted criteria or strategies. Secondly, there is no simple relationship between performance and effort invested. The expended amount of effort depends very much on the structure of the task (data-limited versus resource-limited, Norman & Bobrow, 1975) and, related to this, the amount of practice and experience, and of the operator’s state.

G.Mulder (1980) has linked mental workload to a ‘controlled mode’ of information processing. A distinction between two modes of information processing has been proposed (see Schneider & Shiffrin, 1977, Shiffrin & Schneider, 1977): automatic versus controlled information processing. Automatic processing is fast, not conscious, rigid, requires almost no resources or attention and can be performed in parallel. Automation follows frequent, consistent practice. Controlled processing is effortful, serial, conscious, and is flexible. Controlled processing requires the retention of information in working memory, and hence requires resources and attention. According to G.Mulder (1980) the amount of time an operator processes information in this controlled mode is a reflection of mental effort. Also, in general, a task with higher mental demands is expected to lead to a proportional increase in controlled processing time (see also Meijman & O’Hanlon, 1984).

Resource models have traditionally been used extensively in mental workload research (e.g., Gopher & Sanders, 1984) and the framework has proven to be useful in this area of research. However, this does not mean that the multiple-resource model is universally
supported. Kantowitz (1987), not an opponent of capacity theory, has criticized its multidimensionality. Kantowitz considers the theory “too powerful and too difficult to reject” and states “I do not trust a model that cannot be falsified” (p. 91). He suggests that it is too easy to add another resource (‘pool of capacity’) if data do not fit the theory and he draws attention to a hybrid model launched previously by himself and Knight (Kantowitz & Knight, 1976). In that model a single pool of capacity is divided between perceptual and response stages of information processing by a Static Capacity Allocator. However, this model does not pay attention to interference within versus interference between modalities, a very useful aspect of the multiple-resource theory in mental-workload research. Here the multiple-resource theory has functional utility in predicting interference between tasks.

Clearly, the assessment of workload is coupled with task difficulty as experienced by the operator (Gopher & Donchin, 1986), in particular because several reactions to the task demands are possible. Operators can adapt their behaviour and cope with an increase in demand. They can also change their strategy and task goals and accept a lower performance level or they can give up completely (see Meijman & O’Hanlon, 1984). Strategies will also differ between individuals, and some strategies will be more effective and require less effort to reach the same level of performance. In the case of coping with the demand, an increase in effort is exerted while performance remains at the same level. In that case, performance measures will not reflect any change and be insensitive to the increase in workload, while other measures, such as self-report ratings or physiological measures, may well give an indication of effort exerted. In other conditions in which a change in strategy or ‘quitting’ behaviour occurs, measures of effort may remain unchanged or even show a decrease, while performance measures will indicate decreased task performance.

Terminology in mental workload research has its roots in cognitive and physiological theories. As a result, the terms used are sometimes unclear, as different authors use the same terms with differing meanings. In this thesis task demands, workload and effort are prime concepts. Task demands are determined by goals that have to be reached by performance. These goals can be defined in general terms such as ‘the aircraft should land safely’. It is important to acknowledge that sub-goals are quite often self-set, e.g., first action A then B (or the other way around), and that giving priority to sub-goals can influence general goals and demand. Workload is the result of reaction to demand; it is the proportion of the capacity that is allocated for task performance. Effort is a voluntary mobilisation process of resources. State-related effort is exerted to maintain an optimal state for task performance while task-related effort is exerted in the case of controlled information processing.
Driver workload

A model of the main task of the driver is useful in mental workload research in driving. Parkes (1991) defines the primary task of the driver as “safe control of the vehicle within the traffic environment”. As stated in the introduction, car driving is a dynamic control activity in a continuously changing environment. The driving task is not only influenced by the drivers themselves, but also by the behaviour of other traffic participants. It is not an easy task to model driver behaviour. However, a useful model of driving that takes more into account than just ‘safe control’ (Parkes, 1991) has been offered by Michon (1971, 1985) and Janssen (1979). In this model, car driving is described as a complex task with processes at a minimum of three hierarchical levels. At the top level, the strategic level, strategic decisions are made, such as the choice of means of transport, setting of a route goal, and route-choice while driving. At the intermediate level, the manoeuvring level, reactions to local situations including reactions to the behaviour of other traffic participants, take place. At the lowest level, the control level, the basic vehicle-control processes occur, such as lateral-position control. At this level automatic processes occur, while a level higher controlled processing is required. In particular driver-performance measures can be connected to the three levels. For example, steering-wheel movements reflect performance at the lowest level, car following performance and mirror looking are processes at the manoeuvring level, while errors in route choice reflect performance at the strategic level. Demands at all three levels can exceed capacity, and may result in affected performance, and that includes affected performance at other levels. A student driver cannot yet perform all control-level tasks automatically, and workload with respect to vehicle control is high. This may result in neglect of higher level tasks, such as mirror-checking. In a new traffic environment, e.g. driving in heavy traffic in a city abroad, manoeuvre-level tasks may put high demands on visual and central resources leading to affected performance on the other levels. Demands of monitoring other traffic could be so high that following the signs ‘Antwerpen’ is not possible and a turn is missed. In general performance at a higher level will be affected, although it cannot be excluded that under conditions of high manoeuvre-level demand some drivers will, e.g., shift to a wrong gear.

Sources of driver workload may be found both inside and outside the vehicle. A complex junction that has to be crossed or an important conversation on the car-phone will both increase task demands. Since driving is to a very large extent a visual task, demands on visual and central resources will be highest. However, in the years to come, more in-car technology will be installed in vehicles requiring a raise in allocation of auditory resources. The use of car-phones is already widespread and various new electronic intelligent in-car devices are being developed, the use of which will only increase. While it is
environment demands, e.g., having to merge in heavy traffic, increase workload, while the effects of alcohol, persisting monotony and fatigue increase workload by a reduction in capacity (Schneider et al., 1984, Kantowitz, 1992a, Wierwille & Eggemeier, 1993).

Table 1. Factors affecting workload.

<table>
<thead>
<tr>
<th>Driver State Affecting Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>monotony</td>
<td>fatigue</td>
</tr>
<tr>
<td>sedative drugs</td>
<td>alcohol</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Driver Trait Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>experience</td>
<td>age</td>
</tr>
<tr>
<td>strategy</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Environmental Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>road environment demands</td>
<td>traffic demands</td>
</tr>
<tr>
<td>vehicle ergonomics (RTI)</td>
<td>automation</td>
</tr>
<tr>
<td>feedback</td>
<td></td>
</tr>
</tbody>
</table>
unlikely intentional, these devices will increase driver mental workload and possibly affect behaviour negatively, thus becoming a threat to traffic safety. There is another problem with this boost in development of equipment; collision-avoidance systems, traffic-information systems, driver impairment monitors and navigation systems individually can help drivers, but the combined use can result in overload of their information-processing system (Verwey, 1990). In the GIDS\(^1\) project (Michon, 1993) this problem was recognized and the project proposed to add a scheduling system that plans information presentation (Verwey, 1993a). In scheduling, the driver’s personal limitations should be taken into account. But before tasks can be properly scheduled, the effects on driver workload of tasks in isolation and the effects in combination with other tasks that require simultaneous performance, have to be assessed. A GIDS system needs information about the effect of each individual task on workload, preferably dependent upon local situations, before such a system can decide which task or signal to postpone. Likewise, a road authority might like to know whether the road layout at a specific accident blackspot increases driver mental workload before taking action, or a telecom company may decide to promote their voice-activated dialling car-phone that, in terms of workload, can be more safely combined with the primary task of car driving.

Under certain circumstances it is possible that new in-car technology will have the opposite effect of driver overload, and will lead to monotony in task performance. This could happen, as Kantowitz (1992a) pointed out, if new devices are to actually control the vehicle, similar to flight management systems in aviation (see also Wiener, 1987). At present, driver deactivating situations are mainly confined to monotonous motorway driving. The number of these low-stimulus conditions, in which the driver may become deactivated, may however increase if more functions are taken over by technology. There are scenarios for the future in which vehicle control in terms of steering-wheel movements will also be carried out by an automated system, and the “driver’s” actions will be restricted to strategic level decisions (Hancock & Parasuraman, 1992).

A list of factors that affect driver workload is given in table 1. The table displays both driver state, trait and environmental factors that have an influence on workload. Factors may either increase or decrease mental workload. Automation and the allocation of functions may help the driver, e.g. in conditions where environmental demands are high, but could also turn driving into a task of vigilance. In general, feedback is intended to reduce demand, but sometimes it increases workload by providing additional information that has to be processed. High road-

\(^1\) Generic Intelligent Driver Support