Chapter 1
Perceptual and Cognitive Processes in Human Behavior

Kazumitsu Shinohara

Abstract In this chapter, major models of perception and cognition are reviewed with a focus on human information processing and some related research paradigms. The main focus is on the psychological models of perception and cognition that experimental psychology has developed to explicate the psychological functions of human information processing. By adopting these models, many studies have clarified the roles of the brain activities that are closely related to the components of the models. The models are then important for understanding how we perceive and recognize surrounding environments and how we decide to behave in response to them.

Keywords Skill-Rule-Knowledge based model (SRK model) • Selective attention • Cocktail party phenomenon • Divided attention • Spotlight • Orientation • Useful field of view • Visual search • Feature integration theory • Coherence theory • Attentional resources • Multitasking • Working memory

1.1 Introduction

It is necessary for humans to adaptively behave in a variety of environments. Adaptive behaviors include acquiring information from surrounding environments, activating information stored in memory, processing acquired or activated information, and executing a behavior. Perception and attention are the first stage of adaptive behavior. In this chapter, major models of perception and cognition are reviewed with a focus on human information processing and some related research paradigms.

Obviously, perception and cognition fully depend on a wide variety of brain activities. Many researches have contributed to clarify the relationships between psychological phenomena and brain activities, resulting in a huge amount of neuropsychological findings in this domain. Compared to higher cognitive processes, such as thinking and intending, perceptual and cognitive processes are

K. Shinohara (✉)
Graduate School of Human Sciences, Osaka University, Suita, Osaka, Japan
e-mail: sinohara@hus.osaka-u.ac.jp

M. Kasaki et al. (eds.), *Cognitive Neuroscience Robotics B*,
DOI 10.1007/978-4-431-54598-9_1

simple; it is relatively easy to study the relationships between these psychological processes and brain activities under strictly-controlled experimental conditions.

In this chapter, however, only a few findings about brain activity are discussed. The main focus of the chapter is on the psychological models of perception and cognition that have been developed in experimental psychology. These models conceptually pertain to the psychological function of human information processing, although they do not specify how brain activities are related to psychological processes. Most of these models have been developed on the basis of behavioral and neuropsychological findings. By adopting these models, many studies have clarified the roles of the brain activities that are closely related to the components of the models. It is recommended that the reader refer to related neuropsychological studies if detailed information on neuropsychological evidence for the models is needed.

1.2 Human Information Processing

1.2.1 Automatic and Intentional Control of Behavior

In the study of cognitive psychology, human behavior has been regarded as a sequence of information processing: information processing involves obtaining information from the surroundings, understanding the situation on the basis of information, making a decision about how to act, and executing a behavior.

Many psychological models of human behavior distinguish between two modes of behavior: the automatic mode and the intentional and effortful mode. One of the most famous models is the dual process theory, and it assumes System 1 and System 2 (Stanovich and West 2000). According to Kahneman (2002), System 1 is fast, automatic, effortless, associative and difficult to control or modify, while System 2 is slower, serial, effortful, and deliberately controlled. When we are awake, System 1 and System 2 are corporately working. System 1 continuously generates suggestions for System 2, and System 2 often accepts these suggestions with little or no modification, leading to smooth performance with minimum effort. When System 1 encounters some difficulties, System 2 is called upon for detailed and specific processing (Kahneman 2011).

Whether the mode of behavior is automatic or intentional, there are many psychological processes underneath any given behavior. To understand human information processing more analytically, it is necessary to use the model of human behavior that divides psychological processes into specific components, and to design a study that can pinpoint the function of each component in an adequate research method. In this section, a model frequently used in the human factors research, Rasmussen's Skill-Rule-Knowledge based model (SRK model; Rasmussen 1983, 1986, 1987), is described (Fig. 1.1). The SRK model is mainly used for analyzing and classifying human errors. In the human factors research,

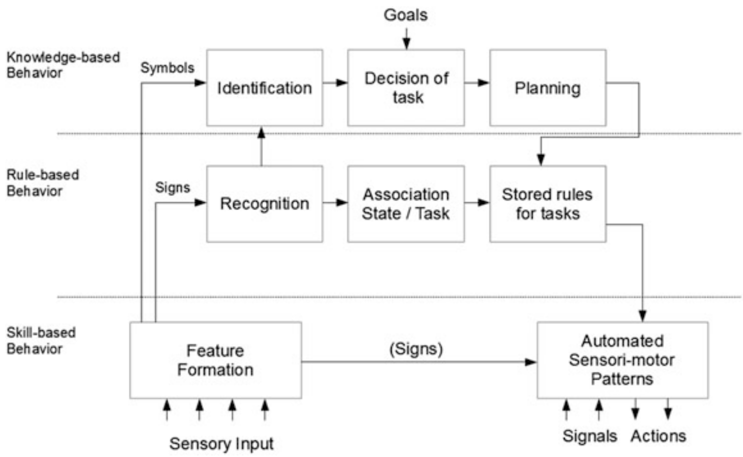


Fig. 1.1 Skill-Rule-Knowledge based model of human information processing (Rasmussen 1983, 1986, 1987)

human behaviors in industrial settings, e.g., factory and machine operation, have been studied for the purposes of accident and error prevention.

In the SRK model, human information processing depends on three levels of cognitive control: the skill-based level, the rule-based level, and the knowledge-based level. Several stages of information processing are assumed at each level.

At the *skill-based* level, a behavior is based on sensori-motor control. Through experience and extensive training, the link between specific stimulus and action is learned. Once this linkage is established, perception of a stimulus immediately activates the action linked to it. It is automatic and not demanding. For example, the performance of an expert sports player is regarded to be processed at this level. An action conducted this way is a pattern of behavior which has been acquainted through extensive experience. Our daily life consists of innumerable acts developed through repeated experience. Basically, a behavior performed at this level does not include intentional control or monitoring, and so modification of ongoing behavior is quite difficult.

At the *rule-based* level, a behavior is controlled by stored rules learned through previous experience. Most of our daily routines are at the rule-based level. After sensory input is recognized, what task is to be performed in the current situation is identified by reference to the stored rules that are appropriate for the current situation. Sometimes the rules are so familiar that the detail of them cannot be consciously specified. At this level, recognized sensory information is used as a sign not for consciously controlling and monitoring the processes but for selecting and modifying the rules to execute the familiar processes.

In the novel, unfamiliar situation in which no rules are available, a behavior is controlled at the *knowledge-based* level. Information processing at this level is consciously controlled, and information is processed as a symbol that refers to a concept to identify the situation. What task is required to perform in the current situation is decided by reference to the explicitly recognized goal, and a sequence of actions is planned accordingly. Rasmussen (1986, 1987) elaborates this decision process at the knowledge-based level and suggests the step ladder model containing eight stages of decision making: activation, observation, identification, interpretation, evaluation, task definition, procedure formulation, and execution. In the step ladder model, data processing in a stage produces states of knowledge, and they are used for data processing in the next stage. Stereotyped mental processes sometimes take a shortcut to the state that is otherwise the result gained by several steps further, and this shortcut may result in an error in decision making. For example, when we observe the situation to obtain information for decision making, we are likely to determine what to do by improperly adopting some habitual rule without checking its pertinence by carefully analyzing the information.

Norman and Shallice (1986) propose a general account of the control of action. In this model, human behavior is controlled in two ways: automatic schema-based control and conscious control by the supervisory attentional system. A schema refers to organized packets of information about the situation, task, events, etc., which are acquired through experience and instructions. When a behavior is based on well-established skills or habits, perceived stimulus automatically activates certain schemas linked to it, and the schemas lead to appropriate actions. When activities come into conflict, the contention scheduling intervenes to resolve the conflict by adopting simple rules and setting priorities on the activities. The contention scheduling is regarded to work relatively automatically. When the automatic control of behavior is unable to sufficiently respond to the situation because of its novelty or complexity, the supervisory attention system (SAS) intervenes in the streams of the automatic behavior routines. The concept of the supervisory control has been used for envisaging the central executive of working memory system (Baddeley 1986, 1997).

The SRK model and the Norman and Shallice model have similar structures with regard to the explanation of human behavior. The skill-based level is comparable to the control based on automatic schema activation; the rule-based level is comparable to the control intervened by the contention scheduling; and the knowledge-based level is comparable to the control by the SAS. These two models are useful to describe the internal processes of human behavior and especially of human behavior in the real world.

1.2.2 Attention in the Human Information Processing

Attention is an important concept for describing human behavior. It is obvious that we can do many things without attention, but we need to use full attention when we face an unfamiliar or difficult task. The three levels of the SRK model have different relations to attention. Attention is not required at the skill-based level, because the direct link from perceived stimulus to the activation of sensori-motor pattern is completely automated. Attention is only minimally required at the rule-based level. Attention is needed to perform each stage of processing at the knowledge-based level. In any case, attention should be taken account of in order to model human behavior in terms of human information processing.

There are three forms of attention: selective attention, divided attention, and sustained attention (Schmeichel and Baumeister 2010). 'Selective attention' refers to the function of selecting information, i.e., focusing attention on one object and ignoring other aspects in the environment. *The cocktail party phenomenon* (Cherry 1953) is such that we can select one voice to attend in the noisy circumstance, such as a cocktail party. The phenomenon is well known as an effect relating to the auditory selective attention. In the early studies of attention, the locus of selection of auditory message has been extensively studied by adopting a procedure called '*dichotic listening*.' In the dichotic listening procedure, different messages are presented to each ear by headphones. Participants are asked to attend to one of the two messages, and to repeat aloud (shadow) it while ignoring the other. After doing this task, participants are asked to answer questions about the unattended message. The typical result is that participants cannot answer what the unattended message says, but can identify the physical features of it, such as the gender of the voice. Interestingly, some subjects can detect their own name from the unattended messages, suggesting that semantic processing is not totally impossible for the unattended information. To explain these phenomena, several models of auditory selective attention have been developed, such as the filter theory (Broadbent 1958), the attenuation theory (Treisman 1964), and the late selection model (Deutsch and Deutsch 1963). These models are concerned with the problem of when the attended auditory information is selected.

'Divided attention' refers to the function of allocating attention to several sources or streams of information simultaneously. It is necessary for post-selection information processing to invest attention as a mental 'resource' or 'fuel'. Attentional resources are divided and allocated to several streams of information processing.

Wickens and McCarley (2008) propose a simple model of attention (Fig. 1.2). In this model, information from external and internal events is filtered and selected at the first stage of processing. Thus, selective attention is regarded as the process of information filtering to avoid the capacity-limited cognitive processes being overloaded. Filtering is affected by such factors, as expectancy, value, salience, and effort. For example, if information is expected to be important and valuable, it is likely to be selected. These factors pertain to the top-down processing, the process consciously controlled to achieve the task goal. Any stimulus with salient features, such as high luminance and prominent color, is easy to be detected, because it automatically captures attention.

Selected information is stored and processed in the working memory. Attentional resources are needed if information processing is to be intentionally executed. Usually, there are several ongoing cognitive processes working at the same time, and so attentional resources are divided and separately supplied to them.

The role of attention can be formulated in the SRK framework. Selective attention is mainly related to the feature formation stage and the recognition stage. Divided attention contributes to the stages of knowledge-based level

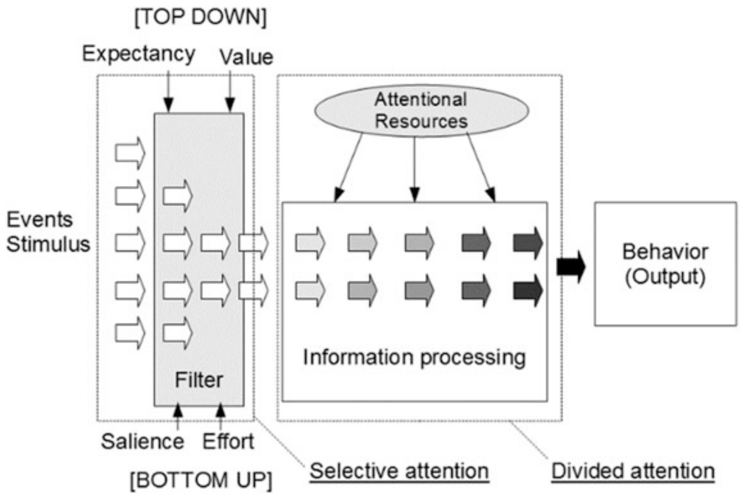


Fig. 1.2 A simple model of attention (Wickens and McCarley 2008)

processes, and to those of rule-based level processes to a lesser extent. The skill-based level processing is usually unrelated to attention control; but if the automatic activation of action is to be inhibited, divided attention is used to interrupt the link between sensory input and action (Shinohara 2011).

1.3 Perceptual Processes

Perception is "the acquisition and processing of sensory information in order to see, hear, taste, or feel objects in the world and it also guides an organism's actions with respect to those objects" (Sekuler and Blake 2001). There are numerous important findings and topics on human perception, such as illusion, depth and size perception, perceptual organization, and so on. For lack of space, this section only reviews the acquisition and selection of visual information as an important perceptual process. The concept of *visual focused attention* is introduced in order to describe this process.

1.3.1 Orientation of Visual Attention

The popular metaphor that visual focused attention is like a *spotlight* has been used for describing the character of visual attention (Posner et al. 1980). This metaphor assumes that information in the spatial area on which the spotlight is casted is rapidly selected and efficiently processed, and that attention can move as the

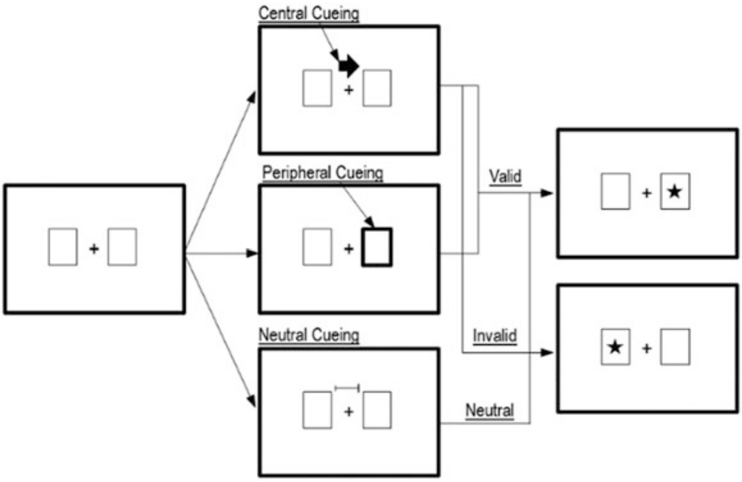


Fig. 1.3 Event sequence in the spatial cueing paradigm

direction of spotlight changes. To study the movement of attentional spotlight, or in other words, to study the characteristics of the orientation of visual attention, the spatial cueing paradigm has been adopted (Posner 1980; Posner et al. 1980).

In the typical spatial cuing paradigm (Fig. 1.3), participants are required to fix their eyes on a fixation point, and to make a response when they detect the onset of the target stimulus. Before presenting the target, a cue pointing to the position of the target is provided. The validity of the cue is systematically manipulated; a *valid* cue actually indicates the position of the target and an *invalid* cue does not. The cue does not always indicate a specific location, in which case the cue is called *neutral*. Typically, the valid cue is presented in the majority of trials. Participants usually move their attention to the position that the cue indicates. When the cue is invalid, participants have to shift their attention to the position where the target is actually presented. Since the presentation of the cue and target is only for a short time and participants are instructed not to move their eye during the trial, the orientation of visual attention in the spatial cueing paradigm is covert.

There are two types of cues: central cues and peripheral cues. A central cue is presented near the fixation point in the form of a symbol or character, and semantically indicates the position at which the target is supposed to appear. Participants are expected to actively move their attention according to the central cue. A peripheral cue is presented at the position of the target. A peripheral cue indicates the position of the target physically and directly. For example, a brief illumination of a possible location of the target or a changing color of the placeholder in which the target is subsequently presented has been used as a peripheral cue. It is expected that visual attention is automatically captured by the onset of a peripheral cue. The shifts of attention induced by a central cue and a peripheral cue are called the ‘endogenous orientation’ and the ‘exogenous orientation,’ respectively (Jonides 1981).

Cost-benefit analysis is used to analyze the result of a spatial cuing paradigm: the cost is the difference in reaction time between the invalid condition and the neutral condition, and the benefit is the difference in reaction time between the valid condition and the neutral condition. Thus, the cost reflects how long it takes to move attention from the cued position to the un-cued position, and the benefit reflects how much time is saved in attending to the cued position. Not surprisingly, the typical result is that reaction is faster and more accurate in valid trials than in invalid trials. When the cue-target interval is short, the benefit of valid cue is obtained for the exogenous orientation, but not obtained for the endogenous orientation. This indicates that the exogenous orientation by a peripheral cue is rapid and automatic (Müller and Rabbitt 1989).

Though the exogenous orientation of visual attention seems to be a simple and transient perceptual process, it actually involves a coordinated operation of several brain areas. Posner et al. (Posner and Cohen 1984) conducted an experiment using a spatial cuing task for neglect patients, and found that the effect of cue validity was different among participants with different areas of lesion. By analyzing the pattern of the results, Posner et al. suggested that there were several processes of attentional control, and that the activities of the posterior parietal lobe, the superior colliculus, and the pulvinar nucleus are respectively responsible for disengaging attention from the previously attended position, shifting attention from the disengaged position to the newly oriented stimulus, and engaging attention to stimulus (Posner and Peterson 1990).

1.3.2 Spatial Width of Visual Attention

The spotlight metaphor of visual attention mentioned above assumes that visual attention is distributed around the fixation point, and that it can be moved to the location where detailed visual information can be acquired. Some researchers have emphasized that the spatial area over which visual attention is distributed is variable according to the task requirement.

While the spotlight metaphor is useful to describe the characteristics of the orientation of visual attention, the zoom lens metaphor (Eriksen and St. James 1986) is proposed to describe the variable size of the area of visual attention. For example, LaBerge (1983) presents a five-letter word and asks participants one of the two questions: whether or not the middle letter in a word is a target and whether or not the word is a name (Fig. 1.4). The former instruction leads participants to focus their attention on the middle letter, and the latter instruction does to focus their attention on the whole word. Additionally, some trials present a row of #s with a probe letter at each letter of a word, and asks participants to respond to the probe. When participants focus on the middle letter in a word, the response to the probe appeared in the position of the middle letter is fastest, and the response to the probe farthest from the center is slowest. When attention is directed to the whole word, the speed of response to any position is constant. These results suggest that the area of visual attention varies with the task requirement. Eriksen and St. James (1986) examine how the deployment of visual

Fig. 1.4 Hypothesized area of focused attention in LaBerge (1983)

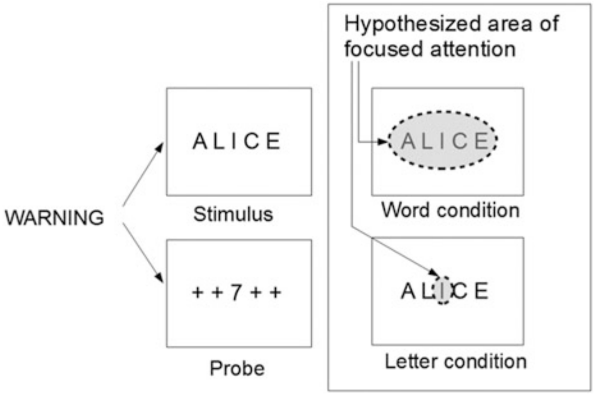
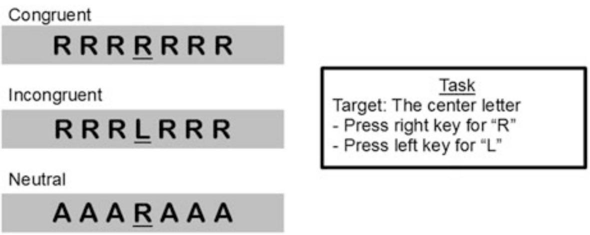


Fig. 1.5 The flanker compatibility paradigm



attention is controlled. In their study, what position is to be attended is indicated by providing cues immediately before presenting stimulus. They have found that the reaction time to stimulus depends on the number of cues and the time interval between cue presentation and stimulus onset. This result suggests that it takes time to manipulate the size of visual attention.

The flanker compatibility paradigm (Eriksen and Eriksen 1974) is useful to measure the size of visual attention distributed around the fixation point. In this paradigm (Fig. 1.5), several letters are presented in line, and each participant is asked to respond to the target letter, the one located in the middle, by pressing a certain key, and to ignore the flanker letters, the ones flanking the target letter. Target and flanker letters may be the same, and only distinguished by location. Responses in the congruent trial in which both the target and the flanker are assigned to the same key are usually much faster than in the incongruent trial in which the response to the target is different from that to the flanker. This phenomenon is called the 'flanker-compatibility effect.' This effect is thought to be caused by the conflict of response information. Both the target and the flanker automatically activate a response based on the pre-determined response mapping. When the target and the flanker activate different responses, the response to the target interferes with the response to the flanker, and a response conflict arises. It is known that the spatial separation between the target and the flanker has a strong influence on the flanker compatibility effect; as the distance between the target and the flanker increases, the flanker compatibility effect decreases (Eriksen and Eriksen 1974; Miller 1991). This is evidence that information in the area of visual attention distributed around the target is inevitably selected and processed. Interestingly, a slight spatial separation, 1 degree of visual angle, can reduce the flanker compatibility effect, suggesting that it can reflect the minimum size of the area of visual attention.

It is important for behavior in the real world to coordinately control the movements of fixation points and the varying field of visual attention around each fixation point for optimal information acquisition from the surrounding situation. Optimized visual information acquisition is a kind of cognitive skill which is obtained through experience. Thus, the pattern of fixation and the width of visual attention have been often examined to analyse the behavior in the real world.

The concept of the useful (functional) field of view, which is similar to the spotlight and the zoom lens metaphor, has been widely used for analyzing visual information acquisition in a real world behavior, such as car driving (e.g., Clay et al. 2005). The definition of the useful field of view is that "the area around the fixation point from which information is briefly stored and read out during a visual task" (Mackworth 1965), or that "the area around the fixation point from which usable information for the recognition of the whole picture is extracted" (Saida and Ikeda 1979). The width of the useful field of view ranges from approximately 4–20 degrees, and varies with many mental and environmental factors, such as the load of central vision, the spatial density and similarity of background objects, arousal level, fatigue, etc. (Miura 2012).

The size of the useful field of view is often critical for real-world tasks. For example, driving a car depends on visual information, and requires the driver to obtain visual information by continuously moving the fixation point around the scene. In essence, it is desirable that the driver can obtain as much information as possible from the area around the fixation point. In other words, if the useful field of view is wide, it is better for safe driving. Miura (1986, 2012) measured the useful field of view of drivers in the real-world setting by using the dual task paradigm. He required participants to drive a car while detecting and responding by pressing a key to the visual stimulus which was occasional emission of LEDs attached on the front window. The size of the useful field of view was estimated on the basis of the reaction time to the visual stimulus. He found that drivers adaptively modulated the size of the useful field of view as the need arose in the driving situation. The movement of visual fixation is minimum, and the size of the useful field of view is large in the low demand conditions, such as highway driving; whereas drivers frequently move their fixation point, and the size of useful field of view is small, in the high demand conditions, such as driving in a congested downtown area.

1.3.3 Visual Search and Feature Integration Theory

In the real world, we often have to find a particular object among many objects by vision. We move our eyes and sequentially check each visual object. This is a visual search task. It has been frequently used as an experimental task to investigate visual attention. A spatial cuing task is used to examine the covert orientation, the movement of attention without eye movement. A visual search task is used to examine the overt orientation, the movement of attention with eye movement.

Typically, a visual search task requires searching for a target defined by physical features (e.g., white circle) in a display with distractors. Participants judge whether the target is in a display or not, and respond by pressing a response key as immediately as possible. How long participants take to detect the target is analyzed. When the target is defined by simple features, e.g., the target is a red circle and the distractor is a blue circle, search time does not depend on the number of visual objects in the display. This indicates that the feature of the target automatically stands out and no search is needed. On the contrary, when the target is defined by a conjunction of several features, e.g., the target is a red circle and the distractor is either a circle or a rectangle and colored either red or blue, search time depends on the number of visual objects. Thus, search time increases as a function of the number of visual objects. The slope of the search time function indicates the efficiency of visual search.

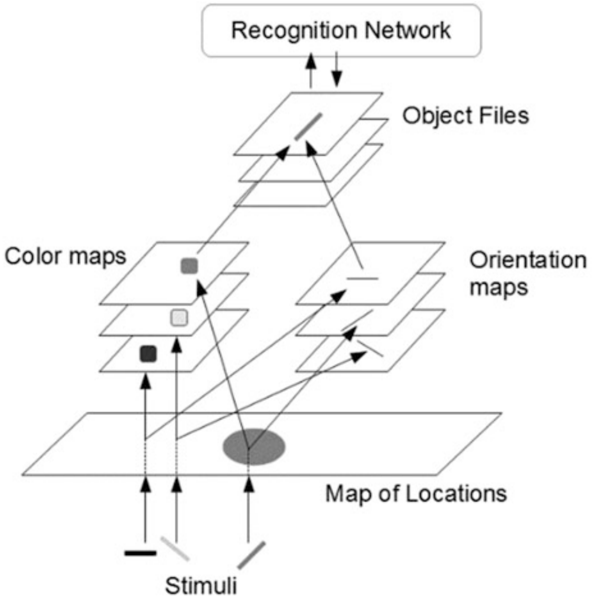
Treisman and Gelade (1980) propose the feature integration theory (Fig. 1.6) of perception of objects. In this model, when visual information is in a particular location of the retina at which the spotlight of attention is directed, basic features of visual information are automatically processed on the basis of separate feature maps, such as color map or orientation map, and then these features of information are combined to create an object. Attention is required to combine features into an object. An object is a temporal representation of visual target, and it is further processed by a higher-order cognitive mechanism, such as recognition network. The target can be found in the feature search without any attentional process; the defining features of the target are processed in a parallel manner. Each visual object must be sequentially searched in the conjunction search; the target can be perceived after its features are combined by the attention demanding process. This process of combining features of a target is called 'feature binding'.

As the primary support for the feature integration theory, Treisman and Schmidt (1982) reports the phenomenon of illusory conjunctions which is a binding error occurred when attention is diverted from the visual display containing several objects. Several studies (e.g. Nakayama and Silverman 1986; Duncan and Humphreys 1989) have reported the findings that cannot be explained by the original feature integration theory, and the feature integration theory has been revised (Treisman and Sato 1990; Treisman 1993).

1.3.4 Scene Perception and Coherence Theory

When an original image A and a partially modified image B repeatedly alternate with a brief blank field between the successive images, it is surprisingly difficult to identify where the modified part is. This phenomenon occurs even when the modification is quite large or when viewers expect that something is different between the images.

Fig. 1.6 Feature integration theory



This phenomenon is called ‘change blindness’ (Simons and Levin 1997; Simons and Ambinder 2005). O’Reagan et al. (1999) report that change blindness occurs even without a blank field; when a brief visual disruption, e.g., an image looking like a “mudsplash,” is presented at the moments of switching images, observers fail to notice differences between them. This suggests that change blindness is related not to masking or occlusion but to selection and representation of visual information. Lavie (2006) conducts a neuroimaging experiment to examine the brain activity during the change blindness task. He has found that the fusiform gyrus, the bilateral parietal lobe, and the prefrontal cortex are active. This finding suggests that frontoparietal activity has a role in visual awareness.

Coherence theory (Rensink 2000, 2002) is able to explain the perception of change (Fig. 1.7a). In this model, incoming visual stimuli are continuously processed, and proto-objects are created. Proto-objects are volatile and last only for a short time. They are easily replaced by new stimuli at their locations. Focused attention selects several proto-objects, and they set up a coherence field, i.e., a reciprocal connection between proto-objects and a single higher level nexus. A higher level nexus pools information contained in an object by summing all inputs.

As for change blindness, to judge which part (object) in the image changes or differs requires object perception. It is difficult to notice a change in any proto-object. When attention is released from a particular location in the image due to a brief blank field or a visual disruption between the images, a coherence field disappears and an object divides into volatile proto-objects.

Rensink (2000) proposes the concept of a virtual representation, according to which a coherent representation of items required for performing the current task is

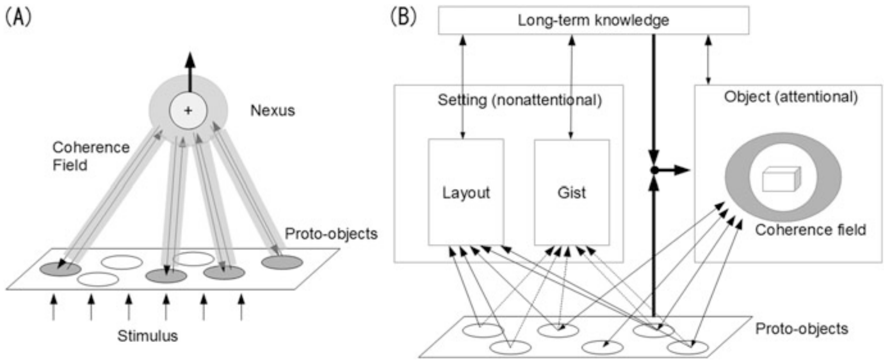


Fig. 1.7 Coherence theory (**a**) and Triadic architecture (**b**) (Rensink 2000, 2011)

created in seeing a scene. It is important that a representation of all items in a scene is not created and coherent representations are created in a just-in-time manner; in other words, they are created when they are needed. Rensink (2000, 2002, 2011) proposes the triadic architecture (Fig. 1.7b) comprising three independent systems: the early visual system, the object system, and the setting system. The early visual system involves an automatic process of creating proto-objects; the object system involves an attentional process to construct coherent fields by linking proto-objects to a nexus; and the setting system is a non-attentional system responsible for the process of extraction of the abstract meaning (gist) and the spatial arrangement of objects in the scene (layout information). Both the attentional "object" system and the non-attentional setting system are linked to the long-term memory, i.e., the memory which activates stored knowledge of the relevant objects and scenes. The former system is executed in a top-down fashion: the access to the long-term memory is intentionally controlled on the basis of the meaning and the importance of each object and scene. The latter system is executed in a bottom-up fashion: it is automatically and compulsory triggered by salient stimuli in the scene.

1.4 Cognitive Processes

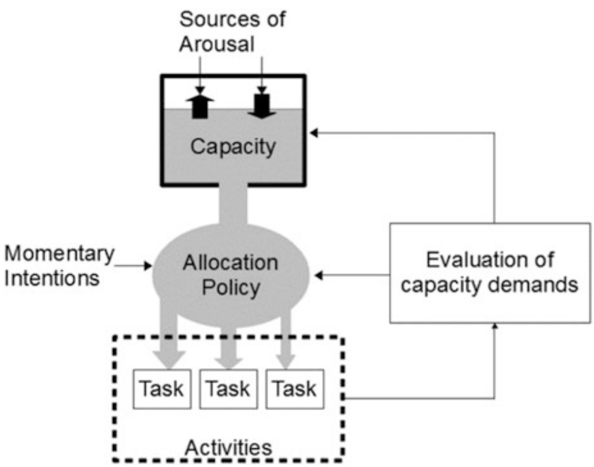
The information selected by the perceptual processes is further processed in the working memory system. It is necessary for information to be consciously processed that attention is allocated to the process at work. The function of attention in the perceptual processes is to select the necessary information to perform the task out of a large amount of information coming from the surrounding environment. The function of attention in the cognitive processes is to maintain the processes for the selected information and to inhibit the processes automatically activated by external stimuli or by internal signals retrieved from the long term memory or schema.

1.4.1 Attentional Resources and Multitasking

When we have multiple tasks to do, it is necessary to divide our attention among them. Especially when tasks are difficult and/or at high stakes, it is crucially important how to allocate attention. Some people may believe that they have an excellent capacity to multitask. It has been reported that some people, who are called "super taskers," show excellent multitasking performances (Watson and Strayer 2010). However, the ability to multitask is in principle limited. For example, when Charron and Koechilin (2010) examined the role of the medial frontal cortex (MFC), the brain region involved in motivating and selecting behavior, both in the single task and in the dual task, they found that the right and left MFCs processed two separate task goals concurrently, and the anterior prefrontal cortex (APC) coordinated these two processes. This suggests that the multitasking ability is limited and can pursue only two concurrent goals. Therefore, even if one seems to simultaneously perform many tasks, it is actually based on the management of two cognitive processes in a time sharing manner. As for divided attention in multitasking, allocation of attention has two purposes: to maintain the processes for each task and to control the time sharing process.

Kahneman (1973) proposes the *unitary-resource theory* according to which common attentional resource, which is closely related to physiological arousal, can be used to sustain a wide variety of task performances (Fig. 1.8). Attentional resources, albeit limited, are allocated to several tasks as they are necessary; allocation of attentional resources depends on the allocation policy. When several tasks are performed simultaneously, the performance of each task depends on the amount of resources each task demands and the amount of supplies from the general resource pool. Capacity demands are evaluated after tasks are performed, and then the allocation policy is modified. The amount of resources can vary as the arousal level changes; the amount of mental resources needed for a performance may be

Fig. 1.8 Unitary-resource model (Kahneman 1973)



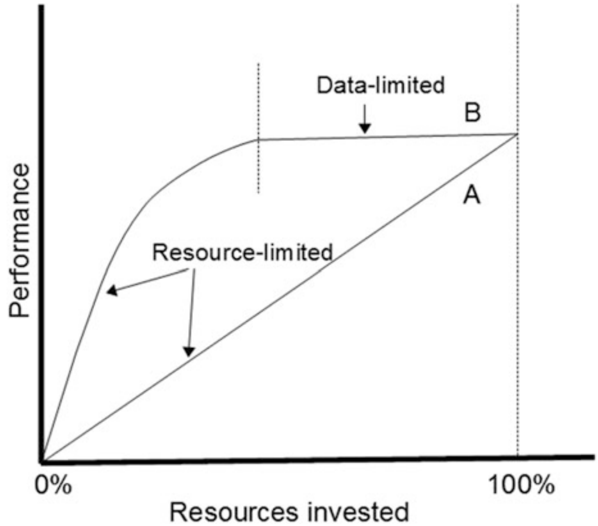


Fig. 1.9 The performance resource function

explained by the difficulty of the performed task (Wickens and McCarley 2008). As the Yerkes Dodson Law (Yerkes and Dodson 1908) indicates, task performance is not a simply-increasing linear function of the amount of available attentional resources. The optimal amount of attentional resources for a task is presumed to be determined by the characteristics of the task, such as its difficulty and complexity.

Task performance usually increases with the amount of attentional resources allocated to the task. However, task performance is affected by the characteristics of information processing during the task, as well as attentional resources. To describe the relationship between task performance and attentional resources, the performance resource function (Norman and Bobrow 1975) has been used (Fig. 1.9). If the task is difficult and must be performed intentionally, the task performance linearly improves as the amount of attentional resource increases (curve A). The task with this feature is *resource-limited*. On the contrary, if the task is easy, well automated to some extent, or if perceptual and/or memorized information for the task is sufficient, the task performance can be maximal with a little amount of attentional resources; there is no improvement in performance if more attention is allocated (curve B). The task with this feature is *data-limited*.

Though the general resource theory is useful to explain many phenomena concerning attention and performance, it cannot explain certain aspects of task performance. When we drive a car while listening to music from the radio, it causes no problem in most cases. This example suggests that different kinds of attention are involved in visual task and auditory task performances.

Wickens proposes the *multiple resource theory* (Fig. 1.10) (Wickens 1984, Wickens and Hollands 1999, Wickens and McCarley 2008). In this model, there are five dimensions of attention resources: the stages of processing, modalities, processing codes, responses, and visual processing. An important assumption of the

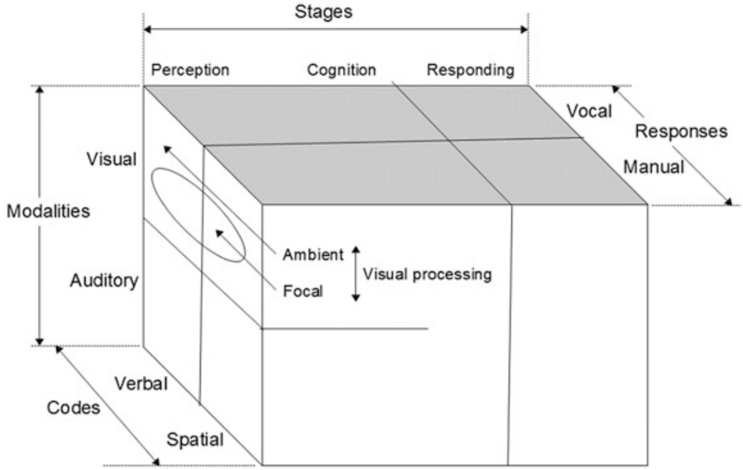


Fig. 1.10 Multiple resources model (Wickens 1984, Wickens and Hollands 1999, Wickens and McCarley 2008)

multiple resource view is that task performance is not disrupted when different attention-demanding tasks depend on different dimensions of attention.

The "stages of processing" dimension involves perceptual and cognitive processing, and selection and execution of response. It has been found in dual task studies that manipulating the difficulty of response does not affect the concurrent performance of perceptual and cognitive tasks (e.g., Wickens and Kessel 1980).

The "modalities" dimension is intuitively easy to understand with an example. It is easy in most situations to read a book while listening to music; this is so especially when the book is readable and the music is relaxing. Studies of human-machine interface have shown that cross-modal displays are better than intramodal displays (e.g., Wickens et al. 1983).

In addition, focal and ambient vision are distinguished with regard to visual processing channels. Wickens and Hollands (1999) suggests that focal vision is used for fine detail and pattern recognition, and ambient vision involving peripheral vision is used for sensing orientation and ego motion. These two aspects of visual processing have different roles in acquiring visual information, and contribute to efficient time-sharing among concurrent visual tasks. For example, when we drive a car, we obtain information about other cars, road signs and signals, by focal vision. At the same time, we use ambient vision to obtain information about the location, speed, and direction of the car we drive.

The "processing codes" dimension is related to the form of information coded in the working memory. As is previously discussed, there are storages for verbal, categorical, or symbolic information, and for analog and spatial information in the working memory system. The separation of spatial and verbal resources is demonstrated by the classic study of Brooks (1968). In his study, participants were asked to visually imagine a block capital letter in their mind and navigate through it. Then, they were asked to fixate their "mind's eye" at the bottom left of the letter and examine whether each corner involved the bottom or top line of the letter in a clockwise, one-by-one manner. Participants answered "Yes" if the corner was on the bottom or top line and "No" if it was not, either by verbal response or by pointing to the word on the paper. The result showed that it is far more difficult to respond by pointing than by verbal response, suggesting that the verbal codes for verbal response, not the spatial codes for pointing, interfere with the visuospatial codes dealing with the letter.

The processing codes dimension is closely related to the "responses" dimension. Wickens et al. have found that a manual tracking task and a discrete verbal task are performed simultaneously with minimal interference. This finding suggests that manual responses, including tracking and pressing, are usually spatial (Wickens et al. 1983, Wickens and Liu 1988).

The dual-task paradigm has been extensively used in the psychological study of attention. In the dual task paradigm, participants are required to perform two tasks simultaneously or each task singly. Manipulating experimental factors, such as combination of tasks, task difficulty, and task priority, task performance in the dual task condition is compared with task performance in the single task (baseline) condition. In many cases, and particularly in the cases of real world task, primary and secondary tasks are distinguished: the former has high priority and is more important than the latter. Assuming with the general resource view that common attentional resources are shared between tasks and the primary task is given priority over the secondary task, the performance of the secondary task is likely to be affected by experimental manipulation. Thus, the performance of the secondary task is used as an index of residual attention which is not allocated to any cognitive processes and is still available for additional allocation. This unused attentional resource is called 'spare' attentional resources. This type of dual task paradigm is called the 'subsidiary task paradigm.'

Given the multiple resource view, the amount of interference in performance of combined two tasks can be examined. If no interference is observed, it is inferred that two tasks do not share the common dimension of attentional resources.

1.4.2 Working Memory and Attention

The working memory is responsible both for holding information for a short term and for actively processing information by interacting with the long term memory. According to Baddeley (2000), the working memory system has a central executive system and sub-systems: the latter include a phonological loop for holding speech-based information, a visuospatial sketchpad for holding spatial and visual coded information, and an episodic buffer for holding and integrating a variety of information. There are interactions between each sub-system and its corresponding function in the long term memory. The function of the working memory has been studied in psychological experiments, and the brain activities underlying each component of the working memory have been extensively investigated in the neuroimaging and neurological studies of patients with brain lesions (see Chap. 3).

The concept of the working memory is closely related to that of attention. Baddeley (1993) states that the general term ‘attention’ is used to refer to the control processes operating throughout the working memory system, and that the label “working attention” would have been used if the working memory system had been studied with a focus on its control mechanisms. Furthermore, Baddeley and Logie (1999) describes the role of the working memory as a mediator of conscious awareness that maintains and coordinates “information from a number of sources including the present, specific episodes from the past, and projections to the future”. This function seems to include attentional processes.

The central executive is regarded as an attentional system, which is a modality-free component operating in the whole cognitive system (Baddeley and Logie 1999). It has endogenous functions of receiving and processing information from the external world to adjust internal task goals, and of selectively generating actions. Baddeley (1996) proposes four functions of the central executive: focusing attention, dividing attention across different sources, switching attention between tasks, and using attention to link the working memory with the long term memory. As is mentioned before, the supervisory attention system (SAS) proposed by Norman and Shallice (1986) is a possible model for the central executive.

The capacity of the working memory is different from person to person, and typically assessed by the span task. The capacity often affects the performance of attentional tasks. For example, when a dichotic listening task requires attending to the message from one ear and ignoring the message from the other ear, participants with low working memory span notice their name contained in the unattended message more often than participants with high working memory capacity (Conway et al. 2001). This finding seems counter-intuitive at first glance, but it can be reasonably interpreted based on the function of working memory. While participants with high capacity efficiently focus on the message from one ear and ignore the message from the other ear, participants with low capacity inevitably pay attention to the message from the other ear. This is because the latter subjects cannot control attention well, compared to the former subjects.

Neuroimaging studies (e.g. Bunge et al. 2001; D’Esposito et al. 1995) have revealed that the brain areas responsible for the functions of the central executive are the dorsolateral prefrontal cortex (DLPFC) and the anterior cingulate cortex (ACC). The DLPFC is responsible for control attention in the dual task performance, and the ACC is involved in conflict monitoring and response inhibition. Osaka et al. (2004) examine focused attention in the performance of the reading span test with fMRI, and find that the left DLPFC, the ACC, and the left superior parietal lobule (SPL) are activated. They propose that the SPL has a role of visual attention controller, and fixate and shift attention in accordance with the ACC and the DLPFC (Osaka and Osaka 2007).

Baddeley’s model of the working memory can be regarded as a kind of multiple resource model. For example, phonological information and visuospatial information are separately stored in the phonological loop and the visuospatial sketchpad, respectively. This separation between phonological process and visuospatial process is similar to the “processing codes” dimension of Wickens’ multiple resource theory (see Sect 1.3.1).

In the study of the working memory, as well as in the neuropsychological approaches, the dual task paradigm has been used to characterize the usage of the working memory in performing a particular task. A *loading task* imposes a processing load exclusively on a particular component of the working memory to suppress its function. When the component of the working memory essential for the task in question is suppressed by the loading task, the performance of the primary task is expected to be significantly impaired. The articulatory suppression suppresses the phonological loop by requiring reading aloud simple words repeatedly; the spatial tapping suppresses the visuospatial sketchpad by requiring pressing spatially arranged keys in an instructed order; and the random number generation suppresses the central executive by requiring continuously producing digits or characters in a random order. For example, Robbins et al. (1996) examine how chess players use their working memory to select a chess moves in a dual task experiment. Chess players are asked to select moves while performing loading tasks. The experiment reveals that the secondary task supressing the central executive and the visuospatial sketchpad impairs the quality of selected moves; the loading task has little effect on the phonological loop. These results suggest that the cognitive skills of chess depend on the processes involved in the central executive and the spatial sketchpad.

1.5 Summary

In this chapter, several models and concepts of human information processing pertaining to perceptual and cognitive processes have been reviewed. A wide variety of models for human information processing in perception and cognition have been proposed and examined by researchers. New findings have accumulated in many ways: psychological experiments with well-developed tasks and paradigms, neuropsychological studies with functional brain imaging and physiological techniques, and neurological studies of patients with brain damages or decreased brain functions. On the basis of the proposed models, we can understand how we perceive and recognize surrounding environments and how we decide to behave in response to them.

It should be noted that it is important not only to elaborate a model by examining the relationships between the functions of perceptual and cognitive processes and the brain activities, but also to consider human behavior in the real world. Human behavior in the real world depends on real-time, simultaneous processing of sensory, perceptual, and cognitive information. Obviously, it is very difficult to establish a model of human behavior which can comprehensively explain behavior in all sorts of situation. At the first stage of explanation, it is necessary to accumulate observational and case data of human behavior in particular situation, and to interpret the underlying structure of psychological processes by referring to the basic models of perception and cognition. An integrative model of human cognition, which is necessary for robots in the future behaving as well as human, will be attained through accumulating these local models of human behavior. If the model is perfect in the true sense, robots operated by the model may make various mistakes as if humans do.

Exercise

Distracted driving happens when one drives a car while engaging in other activities, such as mobile phone use and texting. It can take driver's attention away from driving and induce traffic accidents. Describe the degradation of driver's abilities while distracted driving on the basis of the psychological models of human perception and cognition discussed in this chapter.

References

Baddeley, A.D.: Working Memory. Oxford University Press, New York (1986)

Baddeley, A.D.: Working memory or working attention? In: Baddeley, A., Weiskrantz, L. (eds.) Attention: Selection, Awareness and Control. A Tribute to Donald Broadbent, pp. 152–170. Clarendon Press University, Oxford (1993)

Baddeley, A.D.: Exploring the central executive. Q. J. Exp. Psychol. **49**, 5–28 (1996)

Baddeley, A.D.: Human Memory: Theory and Practice Revised Edition. Psychology Press, Hove (1997)

Baddeley, A.D.: The episodic buffer: a new component of working memory? Trends Cogn. Sci. **4**, 417–423 (2000)

Baddeley, A.D., Logie, R.H.: Working memory: the multiple-component model. In: Miyake, A., Shah, P. (eds.) Models of Working Memory: Mechanisms of Active Maintenance and Executive Control. Cambridge University Press, New York (1999)

Broadbent, D.: Perception and Communication. Penguin Press, London (1958)

Brooks, L.: Spatial and verbal components of the act of recall. Can. J. Psychol. **22**, 349–368 (1968)

Bunge, S.A., Ochsner, K.N., Desmond, J.E., Glover, G.H., Gabriel, J.D.E.: Prefrontal regions involved in keeping information in and out of mind. Brain **124**, 2072–2086 (2001)

Charron, S., Koechlin, E.: Divided representation of concurrent goals in the human frontal lobes. Science **16**, 360–363 (2010)

Cherry, E.C.: Some experiments on the recognition of speech, with one and with two ears. J. Acoust. Soc. Am. **25**, 975–979 (1953)

Clay, O.J., Wadley, V.G., Edwards, J.D., Roth, D.L., Roenker, D.L., Ball, K.K.: Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: current and future implications. Optom. Vis. Sci. **82**, 724–731 (2005)

Conway, A.R.A., Cowan, N., Bunting, M.F.: The cocktail party phenomenon revisited: the importance of working memory capacity. Psychon. Bull. Rev. **8**, 331–335 (2001)

D'Esposito, M., Detre, J.A., Alsop, D.C., Atlas, R.K., Grossman, M.: The neural basis of the central executive system of working memory. Nature **378**, 279–281 (1995)

Deutsch, J.A., Deutsch, D.: Attention: some theoretical considerations. Psychol. Res. **70**, 80–90 (1963)

Duncan, J., Humphreys, G.W.: Visual search and visual similarity. Psychol. Rev. **96**, 433–458 (1989)

Eriksen, B.A., Eriksen, C.W.: Effects of noise letters upon the identification of a target letter in nonsearch task. Percept. Psychophys. **16**, 143–149 (1974)
Eriksen, C.W., St. James, J.D.: Visual attention within and around the field of focal attention: a zoom lens model. Percept. Psychophys. **40**, 225–240 (1986)
Jonides, J.: Voluntary versus automatic control over the mind's eye's movement. In: Long, J.B., Baddeley, A.D. (eds.) Attention and Performance, IX, pp. 187–203. Lawrence Erlbaum, Hillsdale (1981)
Kahneman, D.: Attention and Effort. Prentice-Hall, Englewood Cliffs (1973)
Kahneman, D.: Thinking, Fast and Slow. Penguin, London (2011)
Kahneman, D.: Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. Nobel Prize Lecture. http://www.nobelprize.org/nobel_prizes/economics/laureates/2002/kahnemann-lecture.pdf (2002)
LaBerge, D.: Spatial extent of attention to letters and words. J. Exp. Psychol. Hum. Percept. Perform. **9**, 371–379 (1983)
Lavie, N.: The role of perceptual load in visual awareness. Brain Res. **1080**, 91–100 (2006)
Mackworth, N.: Visual noise causes tunnel vision. Psychon. Sci. **3**, 67–68 (1965)
Miller, J.: The flanker compatibility effect as a function of visual angle, attentional focus, visual transients and perceptual load: a search for boundary conditions. Percept. Psychophys. **49**, 270–288 (1991)
Miura, T.: Coping with situational demands: a study of eye movements and peripheral vision. In: Gale, A.G., Freeman, M.H., Haslegrave, C.M., Smith, P., Taylor, S.P. (eds.) Vision in Vehicles, pp. 205–216. Elsevier Science Publishers, B.V, Amsterdam (1986)
Miura, T.: Visual behavior: research on eye movement, useful field of view, and depth attention relating to this study. In: Miura, T. (ed.) Visual Attention and Behavior: Bridging the Gap Between Basic and Practical Research. Kazama Shobo, Tokyo (2012)
Müller, H.J., Rabbitt, P.M.A.: Reflexive orienting of visual attention: time course of activation and resistance to interruption. J. Exp. Psychol. Hum. Percept. Perform. **15**, 315–330 (1989)
Nakayama, K., Silverman, G.H.: Serial and parallel processing of visual feature conjunctions. Nature **320**, 264–265 (1986)
Norman, D.A., Bobrow, D.G.: On data-limited and resource-limited processes. Cogn. Psychol. **7**, 44–64 (1975)
Norman, D.A., Shallice, T.: Attention to action: willed and automatic control of behavior. In: Davidson, R.J., Schwarts, G.E., Shapiro, D. (eds.) Consciousness and Self-Regulation. Advances in Research and Theory, vol. 4, pp. 1–18. Plenum Press, New York (1986)
O'Reagan, J.K., Rensink, R.A., Clark, J.J.: Change-blindness as a result of 'mudsplashes'. Nature **398**, 34 (1999)
Osaka, M., Osaka, N.: Neural bases of focusing attention in working memory. In: Osaka, N., Logie, R.H., E'sposito, M. (eds.) The Cognitive Neuroscience of Working Memory, pp. 99–118. Oxford University Press, Oxford (2007)
Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., Shibasaki, H.: The neural basis of executive function in working memory: an fMRI study based on individual differences. Neuroimage **21**, 623–631 (2004)
Posner, M.I.: Orienting of attention. Q. J. Exp. Psychol. **32**, 3–25 (1980)
Posner, M.I., Cohen, Y.: Components of visual orienting. In: Bouma, H., Bouwhuis, D.G. (eds.) Attention and Performance X, pp. 531–556. Erlbaum, Hillsdale (1984)
Posner, M.I., Peterson, S.E.: The attention system of the human brain. Annu. Rev. Neurosci. **13**, 25–42 (1990)
Posner, M.I., Snyder, C.R.R., Davidson, B.J.: Attention and the detection of signals. J. Exp. Psychol. Gen. **109**, 160–174 (1980)
Rasmussen, J.: Skills, rules, and knowledge; signals, signs and simbols, and other distinctions in human performance model. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 257–266 (1983)

Rasmussen, J.: Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering. North Holland, New York (1986)
Rasmussen, J.: The definition of human error and a taxonomy for technical system design. In: Rasmussen, J., Duncan, K., Leplat, J. (eds.) New Technology and Human Error, pp. 23–30. Wiley, Chichester (1987)
Rensink, R.A.: The dynamic representation of scenes. Vis. Cogn. **7**, 17–42 (2000)
Rensink, R.A.: Change detection. Annu. Rev. Psychol. **53**, 245–277 (2002)
Rensink, R.A.: The management of visual attention in graphic displays. In: Roda, C. (ed.) Human Attention in Digital Environment, pp. 63–92. Cambridge University Press, Cambridge (2011)
Robbins, T., Anderson, E., Barker, D., Bradley, A., Fearneyhough, C., Henson, R. …, Baddeley, A.: Working memory in chess. Mem. Cogn. **24**, 83–93 (1996)
Saida, S., Ikeda, M.: Useful visual field size for pattern perception. Percept. Psychophys. **25**, 119–125 (1979)
Schmeichel, B., Baumeister, R.F.: Effortful attention control. In: Bruya, B. (ed.) Effortless Attention: A new Perspective in the Cognitive Science of Attention and Action, pp. 29–49. MIT Press, Cambridge (2010)
Sekuler, R.W., Blake, R.: Perception, 4th edn. McGraw-Hill, New York (2001)
Shinohara, K.: Attention and human error. In: Harada, E., Shinohara, K. (eds.) Theories and Applications of Cognitive Psychology 4: Attention and Safety, pp. 186–208. Kitaoji Syobo, Kyoto (in Japanese) (2011)
Simons, D., Ambinder, M.: Change blindness: theory and consequences. Curr. Dir. Psychol. Sci. **14**, 44–48 (2005)
Simons, D., Levin, D.: Change blindness. Trends Cogn. Sci. **1**, 261–267 (1997)
Stanovich, K.E., West, R.F.: Individual differences in reasoning: implications for the rationality debate. Behav. Brain Sci. **23**, 645–665 (2000)
Treisman, A.: Verbal cues, language, and meaning in selective attention. Am. J. Psychol. **77**, 206–219 (1964)
Treisman, A.: The perception of features and objects. In: Baddeley, A., Weiskrantz, L. (eds.) Attention: Selection, Awareness and Control. A Tribute to Donald Broadbent, pp. 5–35. Clarendon Press University, Oxford (1993)
Treisman, A., Gelade, G.: A feature-integration theory of attention. Cogn. Psychol. **12**, 97–136 (1980)
Treisman, A., Sato, S.: Conjunction search revisited. J. Exp. Psychol. Hum. Percept. Perform. **16**, 459–478 (1990)
Treisman, A., Schmidt, H.: Illusory conjunctions in the perception of objects. Cogn. Psychol. **14**, 107–141 (1982)
Watson, J.M., Strayer, D.L.: Supertaskers: profiles in extraordinary multitasking ability. Psychon. Bull. Rev. **17**, 479–485 (2010)
Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Davies, D.R. (eds.) Varieties of Attention, pp. 63–102. Academic, New York (1984)
Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance, 3rd edn. Prentice Hall, Upper Saddle River (1999)
Wickens, C.D., Kessel, C.: Processing resource demands of failure detection in dynamic systems. J. Exp. Psychol. Hum. Percept. Perform. **6**, 564–577 (1980)
Wickens, C.D., Liu, Y.: Codes and modalities in multiple resources: a success and a qualification. Hum. Factors **30**, 599–616 (1988)
Wickens, C.D., McCarley, J.S.: Applied Attention Theory. CRC Press, Boca Raton (2008)
Wickens, C.D., Sandry, D., Vidulich, M.: Compatibility and resource competition between modalities of input, output, and central processing. Hum. Factors **25**, 227–248 (1983)
Wickes, C.D.: Multiple resources and mental workload. Hum. Factors **50**, 449–455 (2008)
Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. **18**, 459–482 (1908)