Dosher, B. & Lu, Z.L. (2010) Attention:
Theories. In E. Bruce Goldstein (Ed.)
Encyclopedia of Perception, Vol. 1.
Sage Publications, pp. 107-112.

ATTENTION: THEORIES OF

Every moment of our lives, the world bombards us with a multiplicity of sights and sounds and other information taken in through the senses. These stimuli may be relevant or not to your current cognitive goal, or of such potential importance that, unexpected as the input may be, you must break out of your current goal-directed activities to understand the new input. A mother might be looking for her child in the playground, but may still respond to the sound of her cell phone. An air traffic controller may be busy tracking a large number of display inputs while actively interacting with only one or two. These and myriad other examples reveal that we are exquisitely tuned to the ebb and flow of new information. Attention is the mind's solution to the problem of sensory and processing overload. Whether this involves the selection of a particular piece of sensory information for additional processing or carrying out two tasks at once, attention systems select inputs and coordinate processing resources of the human brain. Scientists have furthered their understanding of attention through both empirical and physiological observations and the development of theoretical models for mechanisms of attention, as described in this entry.

Early and Late Selection

Inspired by analogies between humans and computers, the first of the modern theories of attention tried to explain how humans selected some sensory inputs for further processing while ignoring others. These theories also emphasized the extent to which sensory inputs could be processed without attention. *Early selection* theories claimed that the attention system acted like a filter that could select some inputs for additional processing based on visual physical characteristics such as location, color, or texture or based on auditory characteristics such as location, pitch, or the speaker's unique voice. Processing for identification or memory was claimed to occur only for attended-to inputs—unselected inputs were filtered out completely. At the other end of the continuum, *late selection* theories argued that considerable processing could be carried out on many inputs without limitations, and that the bottleneck in processing occurred just before the choice and execution of a response.

Initial tests of these theories were similar in essence to the "cocktail party" phenomenon, where a listener must pick out a single speaker in a room full of competing conversations. Experiments present different messages, sometimes by different speakers, in each ear ("dichotic" listening) using headphones. Listeners might be asked to immediately repeat ("shadow") a target speaker, as the message is heard in one ear, while ignoring messages in the other ear. However, attention filtering was ultimately shown to be by no means complete or absolute. Unattended messages, such as those presented in the other ear, were at least partially processed, especially salient or high-valence messages, as in hearing one's name in the unattended message. People notice if a word in an unattended message is the same or related to the attended message, so long as they occur close together in time. Consistent with this, physiological measures have shown that the effects of attention seem to begin at the first stages of sensory processing. Event-related potentials (ERPs), or electric activity of the brain in response to visual inputs, show stereotypical differences in the responses for attended compared with unattended stimuli. Attention can modify the brain's responses to stimuli, amplifying relevant inputs or attenuating irrelevant ones, even the first and fastest responses. For example, the response over the visual cortex to an attended visual stimulus ("pay attention to targets in the upper right quadrant") causes a larger response within the first 100 milliseconds (ms) in the P100 wave of the ERP, and this extends through early visual responses in cases where attention focuses on one location over longer blocks of many trials. Ignoring inputs does not, however, eradicate the sensory responses. In short, a modified form of the early selection model provides a good account: selection affects the earliest forms of sensory processing, but ignored inputs are attenuated rather than filtered out completely. The magnitude of the difference in response between attended and unattended inputs depends on exactly how attention is controlled. Attention has smaller effects on the earliest brain responses in cases where the focus of attention is changed to a new location by a cue on each trial. So, the impact of attention on perceptual processing can be more profound with a sustained period of attending to one location or feature than if selective attention is shifted moment to moment by changing cues.

Filtering and Amplification Mechanisms of Attention

Early selection theories claimed that attention filters some stimuli out and enhances others, but they failed to do more than sketch the mechanisms of filtering or the consequences for information processing accuracy. However, two related mechanisms of attention have been proposed and tested within computational models of a human observer's performance. In these models, stimuli are processed through perceptual templates tuned to respond to targets. For example, a template might be tuned for the letter A, or a particular shape or sound. The response of the template is noisy, and this and other noisy processes limit the accuracy of identification. Separate mechanisms exist for the enhancement or amplification of an attended stimulus and the filtering out of external noise or distractors. The *template model of attention* specifies how filtering and enhancement separately affect the accuracy of detecting a target or discriminating between targets within a framework of signal detection theory. For visual tasks such as identifying a letter or the orientation of a patch of pattern, tests have been developed that measure identification accuracies in the presence of different amounts of noise in the stimulus, something like looking at a poor television image that has a lot of visual "snow." Here, cueing attention to a location has one widely observed mechanism that filters out the noise—it improves performance most precisely when noise or distractors would otherwise damage performance. A second less frequently seen mechanism, enhancement, improves performance in an attended location even in the absence of competing inputs and is especially important for weak targets. Attention enhancement and attention filtering can occur separately or together. These mechanisms can also be directly related to models of neural processing in the visual cortex.

Signal Detection Theory Framework

Signal detection theory (SDT) explains how performance accuracy depends on the strength of the target ("signal") stimulus compared with nontargets in the context of the variability, or noise, in the signal strength over test trials. SDT assumes a noisy response to each stimulus, and determines how criteria or decision rules classify the stimuli to make a response; it separates the true discriminability of target and distractor stimuli from possible changes in criterion. The SDT framework was the basis of the template model (discussed earlier) and has been important to understanding attention limitations in searching for a target among other items. SDT models make predictions about the accuracy of performance when people look for a target among different numbers of distracting stimuli or locations under conditions where people cannot achieve perfect accuracy.

One interpretation of poor performance with more distractors was that people have a limited capacity for processing many items at once. For example, early behavioral tests of these ideas involved briefly presenting a dim light in one of several locations, and noted that the accuracy of reporting the correct location decreased as the numbers of locations increase. Early researchers took these results at face value to say that it was not possible to process more locations without reductions in performance because of capacity limitations in attention. Today, researchers know that such conclusions require first an analysis of the role of chance guessing (or criteria) and require signal detection computations. Even if each item is processed as well in the presence of distractors as alone, the natural variability in the sensory registration of each item causes more errors with more distractors.

In the signal detection framework for visual search, each location (or item) yields an estimate of sensory strength, or match to a target template (the "signal"). Each estimate, whether from a target or a distractor, is noisy because of processing inefficiencies or variation in the presented stimulus. Each added input to the decision increases the potential for false alarms, or classifying a nontarget as a target in error. Even for an ideal machine in which all evidence is processed perfectly and recalled without capacity limits, added inputs reduce accuracy. Only when the losses in accuracy exceed those predicted for an ideal observer are capacity limits implicated. For almost all displays where a target is defined by a single feature or a simple combination of features, the detection or discrimination accuracy is consistent with the predictions of the unlimited-capacity ideal observer model. This includes simple detection ("did something appear?"), but also discrimination of features such as length, orientation, color, or simple shape ("was there a longer one?"). It also includes many difficult cases of visual search where the target stimulus is defined by conjoining two simple features, such as a given orientation and color, where the distracting nontarget stimuli may share either orientation or color with the target. In all these cases, the increases in errors with the number of display elements have been almost exactly those expected from signal detection computations of increasing false alarms (or reduced hits caused by increasing criteria to avoid those false alarms). Limited capacity seems to be limited to special cases where discriminating a target involves spatial relationships, such as two touching bars in which white is on the left and black on the right, or vice versa. Related models of the speed of decision also support unlimited capacity parallel processes, at least for the accuracy-limited brief displays.

Salience and Guided Search

Guided search or salience models of attention predict visual search time when it is possible to scrutinize displays with eye movements. Models of salience, which mark certain items for priority processing, have illuminated how attention may be guided bottom up from stimulus characteristics and top down through goals. Strong or unique sensory stimuli such as the only vertical item among all horizontal items, or one that appears suddenly, have high salience that makes a search easy if the salient item is the target, or make it more difficult if it is not. Stimulus attributes also can interact with behavioral goals to increase the relevance of some items over others, so if one is searching for a red vertical line, items sharing these features may inherit higher salience. In these models, salience of each item is computed in terms of bottom-up factors such as how different it is from other items, especially items that are close in the visual field. The salience is also determined by top-down factors such as the similarity of each item to attributes of a goal or target. For example, in searching for that red vertical line among red horizontal and green vertical items, the ease of search often depends, all else equal, on whether there is a small subset of red items (or vertical items) to which search can be guided.

Several models of salience-guided search have been developed and tested that provide algorithms for the computation of salience and predictions about how these salience computations interact with distinct behavioral goals and tasks. These salience models, and other forms of guided search model, have modified and improved a set of previous models that were based on the notion that the role of limited-capacity, serial attention processes was to bind back together, or integrate, the features of objects that were analyzed and taken apart by separate feature maps in early visual analysis. The original feature integration models of visual search made a contrast between visual searches without needing attention ("pre-attentive") and those that are sufficiently demanding that they need serial scrutiny. Pre-attentive searches occur when any single feature such as any red item independent of orientation, size, and other features is sufficient to define a target as distinct from all other cases where a conjunction or combination of several features (i.e., red and vertical) is needed to define the target. The search must locate an object that contains both features. However, more refined guided search models account for all the early results of visual search used to support the feature integration theory, but also account for findings that reject the feature integration model, such as the ability to restrict search to the subset of red stimuli. Predicting the time to find a target in a visual search task has been the primary goal of guided search and salience models. The salience computations may also make predictions about the likely focus of eye fixations during visual search. Salience-based attention and search models have also been developed in computer science or machine-vision applications, where image-analysis algorithms or programs that do salience calculations are developed to mimic human-like capabilities (and limitations) in visual search tasks.

Capacity and Scheduling

What happens when people try to carry out multiple tasks at the same time? Successful multitasking could magnify human efficiency. However, as failures such as the dangerous use of cell phones while driving or the inability to benefit from listening to different lectures in each ear show us, there are often important limitations to how much can be simultaneously processed. Models of attention capacity and bottlenecks have been used to explain these limitations in simultaneously carrying out two tasks, each with a decision and response. One approach measures *attention operating characteristics,* graphing performance measures from two tasks against one another while varying the instructed division of effort, from focusing entirely on one or the other and different sharing ratios (e.g., 70 to 30%), to show the extent to which two tasks interfere with one another. For example, accuracy of Task A (e.g., visual contrast increment detection) is graphed against the accuracy of Task B (e.g., an auditory contrast increment detection). Observing no sharing losses, that is, where the two tasks can be performed as well together as separately, corresponds to the absence of mutual capacity limitations, or the "ideal point," where the accuracies of Tasks A and B performed together with equal emphasis ("attend equally" or 50 to 50%) are as high as when each is performed alone. The "switching-line" connects single-task performances (from single Task A performance and chance Task B performance to chance Task A performance and single Task B performance), where performance improvements in one task trade directly with performance losses in the other, reflects choosing one or the other task, but not both, to do on any given trial. Often, performance is worse than the ideal point but better than the switching line, corresponding with some degree of successful sharing. The attention-operating characteristic provides a powerful measure of dual-task compatibility.

In other kinds of tasks, people are asked to carry out one task first, followed as quickly as possible by the other. Presentation of the two stimuli is offset by different amounts of time to measure the *psychological refractory period.* If asked to classify a briefly flashed letter by speaking the name and to classify a tone as high or low by pressing a key, the response to the second-arriving stimulus may be delayed. To be refractory is to be resistant, or unresponsive to a stimulus. If two stimuli occur close together in time, then the response to the second is often delayed as though it were necessary to process the first before there is capacity available to process the second one, often by as much as a second or more. In general, it has been concluded that initial perceptual analysis often is carried out on multiple inputs simultaneously, and the need for limited-capacity attention occurs later, perhaps at the level of decision or response selection. The extent of capacity limitation or interference depends on the demands of the tasks. There are two distinct interpretations. One is that bottlenecks of limited attention are intrinsic to certain stages, such as decision, and the other is that introduction of delays is a strategic choice designed to eliminate conflict at the level of the peripheral systems, such as the eye or the hand, which cannot be in two different places at once, and not a central capacity limit at all. This second theoretical framework has been tested using computer programs that simulate specific tasks and predict the pattern of response times. Whether obligatory or strategic, the existence of temporal delays in responding to two input stimuli close together in time is a general finding demonstrating people's limitations in responding to two tasks.

Control of Attention and Attention Switching

Shifting attention from one location to another is an essential part of how we view the world. Covert attention, movement of the mind's eye, leads and targets the movements of the eye. Covert attention shifts are also used to take up information in different spatial regions without movement of the eye. Instructing people to make shifts of covert attention is one strategy for measuring how attention is shifted from one location to another. In one classic example, an arrow points at and leads attention to a possible target location. The costs of misdirection, called invalid cueing, can be measured in lower accuracy and longer response time compared with valid cueing. In another example, individuals may be asked to search for a target in a stream of letters appearing in one location, one after the next, and then to switch immediately to a second stream to detect the next item. The theoretical metaphor of covert attention operating like a spotlight led some researchers to suppose that moving the mental attention spotlight would entail activating spatial regions between the current location and the new location. Clever theories and experiments, however, have shown instead that items in intermediate locations do not benefit from the shift of attention. Instead, if attention is redirected from one location to another, it is as if one spotlight is dimmed while another spotlight focused at a second location is lighted. Excellent quantitative models account for attention switches as such episodes of information acquisition through attentive deployment that operate separately in space and time.

As in the case of visual search, many of these tasks combine a goal-directed, or top-down, specification of a location or a target with a stimulus-based, or bottom-up, analysis of the items in the field and their salience. Indeed, it seems that many visual searches involve naturally evolving cycles of attention deployment that lead to eye movements, and then to a new reassessment of the search field, and another eye movement. In visual search, in planned sequences of shifts of covert attention, or in unplanned or evolving sequences of movements of the eye, the control and shifting of covert attention plays an important role in how information is acquired from the inputs of the outside world.

Physiological Substrates of Attention

With the explosion of new knowledge about brain structure and function, new physiological theories can describe the impact of attention on individual cells in the sensory analysis systems and in networks of centers that act to engage attention or to disengage and move attention, and to maintain vigilance. Spatial attention changes the strength, or rate, of response for individual neurons in the early visual cortex, especially in the presence of distracting stimuli. Individual neurons may respond as though distracting stimuli have been excluded from influence on the response, a property similar to signal-detection-based template models of attention as a filter. Individual neurons can increase the level of response to an attended stimulus as though the contrast or strength of the stimulus was increased. This property, sometimes called contrast-gain, is directly related to the concept of enhancement or amplification of the attended stimuli in quantitative signal-detection models of attention. Computational model frameworks that integrate these parallel models from the cellular level in one or another individual region of the sensory cortex up through full analysis and behavior remain to be developed. Still, it is clear that the behavioral and the neural responses share close functional analogies.

Distinct networks of brain centers have been implicated in goal-directed preparation for expected inputs or targets and in the reorienting of attention in response to salient or unexpected sensory stimuli. Frontal and parietal regions of the cortex, including frontal eye field and lateral intraparietal areas, are activated during goal-directed preparation or deployment of attention in response to a cue to a relevant location or feature. These systems are activated during the period when the expectations from a cue are maintained and may be recruited during goal-directed tasks such as visual search. A different network, one that is lateralized in the right ventral frontoparietal area, appears to be involved in the resetting and reorientation of attention when an unexpected or salient input overrides goal-directed orienting. In goal-directed situations, recent evidence suggests that frontal activity precedes activity in the relevant sensory areas. If attention is attracted by the stimulus, however, then activity in the sensory representations seems to lead to the frontal activity. These networks have been identified through new imaging technologies that sense the activated regions of the brain during attention tasks in humans, through behavioral anomalies in humans with lesions, and by measuring neural responses in tasks in monkeys that parallel those in humans.

Future of Attention Theories

Attention is a critical function in the mental arsenal, one that has fundamental implications for how we interact with the world. As we learn more about how the brain works, computational models could help link cognitive function to physiological responding, and psychophysical models of function and behavior to the workings of physiological components of the relevant brain networks or systems underlying control of attention. These new theories and models, in turn, would help us better clarify the role of attention. A better understanding of attention's "exquisite control" in determining which inputs are processed and when would further our understanding of intelligent information-seeking behavior and how we choose to manage multiple task demands. In turn, the understanding of the distinct mechanisms and brain circuitry of attention may also increasingly serve as a basis for understanding how the attention systems fail to function or function differently across the life span or in different mental health conditions such as attention deficit disorder or schizophrenia.

Barbara Dosher and Zhong-Lin Lu

See also Attention: Covert; Attention: Divided; Attention: Object-Based; Attention: Selective

Further Readings

Brefcynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience, 2,* 370–374.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science, 315,* 1860–1862.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.

Dosher, B. A. (1998). Models of visual search: Finding a face in the crowd. In S. Sternberg & D. Scarborough (Eds.), *Invitation to cognitive science: Vol. 4. Methods, models, and conceptual issues* (pp. 455–521). Cambridge: MIT Press.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269–297.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55,* 207–234.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes mechanisms of attention. *Vision Research, 38,* 1183–1198.

Pashler, H. (1998). *The psychology of attention.* Cambridge: MIT Press.

Shih, S-I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 109,* 260–305.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences, 24,* 295–300.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford University Press.