

Attention and its Role: Theories and Models

<https://doi.org/10.3991/ijet.v14i12.10185>

Athanasios Drigas (✉), Maria Karyotaki
Institute of Informatics and Telecommunications, Athens, Greece
dr@iit.demokritos.gr

Abstract—Motivation, affect and cognition are interrelated. However, the control of attentional deployment and more specifically, attempting to provide a more complete account of the interactions between the dorsal and ventral processing streams is still a challenge. The interaction between overt and covert attention is particularly important for models concerned with visual search. Further modeling of such interactions can assist to scrutinize many mechanisms, such as saccadic suppression, dynamic remapping of the saliency map and inhibition of return, covert pre-selection of targets for overt saccades and online understanding of complex visual scenes.

Keywords—Attention, consciousness, attentional control, attentional flexibility

1 Introduction

The interplay between bottom-up and top-down influences on overt attention is the key for investigating and modeling attention. More specifically, future research should begin with widening the currently narrow scope on top-down factors. Attentional and emotional processes are in the focus of research, although overall research in this field is just at the beginning and suffers from the diversity of concepts of both attention and emotion. Until now, models of saliency maps and research on elementary image features have not considered emotional components. Thus, it is a challenging task for future research to incorporate bottom-up factors and emotion as a strong top-down factor.

2 Theories of Attention

Chun et al. [5] distinguished attention according to the types of information that attention operates over. Therefore, external attention refers to the selection and modulation of sensory information (perceptual attention), whereas internal attention includes cognitive control as well as the contents of working memory, long-term memory, task sets or response selection (central/reflective attention). Cognitive control mechanisms prioritize, independent of sensory modality, which perceptual information to select, encode and maintain in working memory, while suppressing distraction. Moreover, executive processes and working memory influence perceptual processes and guide

eye movements. However, perceptual attention can also influence what gets maintained in working memory. Unattended objects or events can be processed if the primary target task is easy, indicative of late selection. On the contrary, if the primary task is very difficult, all of attention becomes devoted to the target, revealing patterns of early selection. Increased central load, such as increased working memory load or task switching, increases interference from distractors and results in full perceptual identification of ignored items (late selection). Finally, overt training of attention represents a highly promising area for interdisciplinary and translational research.

Kerr et al. [14] proposed a neural framework to explain how body-focused attention could exert “upward” influence on metacognition and on cognitive and emotion regulation. More specifically, body-focused attentional practice enhances localized attentional control over the 7-14 Hz alpha rhythm that is thought to play a key role in regulating sensory input to sensory neocortex as well as enhancing signal-to-noise properties across the neocortex. The generalization of top-down, attentional alpha rhythm modulation to other thalamocortical circuits could enhance one’s ability to filter and prioritize the flow of information throughout the brain, such as selective attention and working memory. Somatic attentional modulation proved to bring about a more generalized improvement in selective spatial attention both in visual and auditory modalities. Moreover, increased alpha power facilitated working memory processes through suppressing irrelevant, internal stimuli by attending to a sensory stimulus such as breadth. Furthermore, somatotopically focused attention enabled the broader modulation of the sensory field, which in turn enabled a more sustained, yet homeostatically, regulated attention to distressing thoughts, feelings and sensations. Finally, working through the sequence from mindfulness of the body to mindfulness of thoughts evoked one’s ability to maintain greater attentional flexibility and thus, be present, but non-reactive to his internal experiences, such as negative cognitions and strong negative emotions.

Tsuchiya et al. [27] made a review study on the relationship between consciousness and attention. There is a common belief, originating from the pre-frontal parietal network (PPN) that consciousness in different senses is always interrelated with a type of attention. Moreover, the pre-frontal parietal network (PPN) is also associated with working memory, executive control and chunking, thus, attention complements the set of the core psychological components of consciousness. However, others are in favor of attention and consciousness following an independent path in their route to decision making processes. Except from the variable types of attention and consciousness and, possibly, their separable function, researchers point to the intermodal effects of attention on consciousness. Conscious auditory perception, olfactory consciousness as well as conscious retrieval of memories were all found to be influenced by attention. Notwithstanding, there is an ample field for empirical investigation regarding the effects of top-down attention on the conscious perception of an isolated stimulus that is either dominant (visual input or input present in time) or non-dominant (olfaction and memory or past memory/future planning). Therefore, peripheral vision, unexpected strong olfactory stimuli as well as the feeling of familiarity may all interfere with top-down attention and its interrelation with conscious experience.

Chun [4] explains the link between visual working memory capacity and visual attention. According to the limited resource model of attention, the amount of attention allocated to an item is a function of how much overall processing is required in a task. However, visual working memory not only depends on the actively sustained maintenance of relevant sensory representations, but also on the inhibition of distraction. Therefore, working memory encoding and maintenance reflects actively sustained attention to a limited number of visual objects and events; thus in the absence of sensory input, internal attention sustains representations in relevant sensory mechanisms as well as other internal memories, choices and thoughts, while maintaining the binding of features into integrated objects. The aforementioned process resembles a meta-cognitive and conscious process, as the individual is volitionally trained to facilitate his everyday vision and behavior.

Gehring et al. [6] made an experiment with the aim to identify all bottom-up and top-down processes contributing to the switching effect as well as to individuals' top-down attentional control. Researchers implemented a mental counting task measuring switch and no-switch reaction time (RT) difference as an index of the time required to switch attention from one internal counter to another, thus suggesting that the internal focus of attention was limited in capacity. Moreover, the no-switch condition resembled the switch trials as the repetition of the physical features or identity of a stimulus could not deliver faster reaction time (RT). Event-related brain potentials (ERP) were additionally deployed in order to enhance the sensitivity of the experimental measures. Results showed that bottom-up processes included priming and top-down processes included an articulatory rehearsal process, both of which influenced the switching effect. Stimulus repetition as well as the difference between no-switch/different and no-switch same trials, called the stimulus-mismatch effect, and the difference between switch and no-switch/different trials, called the counter-switching effect, impacted on individuals' performance or ERP. In conclusion, internal focus of attention could correspond to the activation and deactivation of counter representations and subvocal articulatory processes, evident in counter priming processes.

Kiyonaga et al. [16] argue that working memory is reciprocally related to attention either voluntarily or not. Attention controls working memory and the representations maintained in working memory can capture visual attention whether internal representations are consistent with the external perceptual goals or not. Moreover, both attentions, internal and external, as well as working memory should be considered as one storage mechanism of prioritization of processing that can be geared at internal or external representations, sharing a common, limited, cognitive resource. Therefore, internal and external attentions closely interact and affect one another in a trade-off relationship within modalities or domains. The dispersion of the internal representations across several items and tasks either favors or hinders perceptual selection. Notwithstanding, external attention demanding tasks into shorter periods of time have an impact on the ability to maintain internally attended information; thus individuals' inability to refresh, rehearse or remove interfering representations from their internal attention representations can alter task goals and bias external selection. On the other hand, just as top-down control of attention can be applied to prioritize a certain stimuli

in a visual or auditory environment, internal attention can be strategically controlled to modify its impact on external selection or external attention deployment task. Future research should investigate individuals' internal representations duration or formation while being exposed to external attention cues.

Philips [22] initiates the concurrent non-perceptual activity during the interval to be timed as a form of internal attention as well as a crucial measure of perceived duration of a cognitive task. He suggests that perceived duration is tied to internal attention as a mental activity within individuals' non-perceptual stream of consciousness. In other words, attention to time is connected reciprocally to the quantity of conscious mental activity with the latter being a key measure of time perception and specifically, perceived duration. The stream of perceptual consciousness is divided by attention and the stream of thought is simply a stream of attention. Therefore, perceived duration is measured by the quantity of concurrent mental activity within a stream of thought and elapsed time may well be considered a change within the stream of thought. Perceived duration judgments can, thus, be measured in units of words, syllables or phonemes in addition to hard tasks having a high burden of unconscious processing, such as manipulating mental arithmetic data or other executive tasks used in dual-task paradigms. Harder tasks demand a lot of unconscious processing and yield fewer changes within the stream of consciousness meaning that harder executive tasks should leave untouched or diminish individuals' estimation of duration. Furthermore, duration perception can be affected by life threatening events, dopamine agonists as well as stimuli preceded or accompanied by a train of repetitive stimulation.

Hollingworth et al. [7] investigated on the qualitatively distinct visual working memory states that interact and do not interact with perceptual selection. Researchers conducted three experiments in order to prove that visual working memory has an active state that interacts with the sensory processing of visual stimuli and an accessory state that does not influence perceptual selection. The first experiment intended to confirm that uncued colors, deprioritized for retention, do not capture attention to matching stimuli and included a color change detection task as well as a visual search task. In the second and third experiment, the probability that a color was retained in visual working memory and the precision of the color memory representation were controlled. The dual-state object representation in visual working memory was verified. Cued colors, prioritized for retention, recruit attention to matching stimuli and influence perceptual selection, whereas uncued colors do not influence perceptual selection. The results of the other two experiments showed that color retention from uncued objects was robust and precise, thus, a top-down control over visual working memory representations may have been developed. Therefore, items maintained in visual working memory interact with perceptual perception through a task-relevant search template, encompassing a top-down attentional set for the guidance of visual search.

Wu Dowd et al. [28] shed light on the qualitatively different states of visual working memory in terms of the varying informational content of the representation and its overlap with perceptual processing. Two types of attentional guidance paradigms were employed to assess memory-based guidance in two visual search tasks; in the binary-stimulus dual-task paradigm, attention is supposed to be captured involuntarily

by memory-matching stimuli, whereas in the unitary-stimulus dual-task paradigm, attention is assumed to be undistracted by memory-matching distractors. In both cases, the stimulus arrangement of the two tasks differed as well as the memory cue never matched the location of the search target. Attentional capture effects in the binary-stimulus and unitary-stimulus paradigms did not correlate. In addition, researchers compared the memory-based capture in each task to the independently measured constructs of visual working memory and self-report measures of attention. Although the amount of variance explained by these factors was relatively small, underlying attentional control processes mediated by memory processes were found to differ on account of the stimulus attributes in each task, such as the spatial concentration of task-relevant features.

Kiyonaga et al. [15] proposed the time-based resource-sharing (TBRS) model of working memory in order to explore the reciprocal relation between visual attention and working memory representations. According to the aforementioned model, a time-shared cognitive resource alternately refreshes internal working memory or selects external information. The internal maintenance of information accruing the individual's attention increases the sensitivity to matching items in the environment. More specifically, more time-intensive demands (number of intervening visual search processes and amount of available time to perform them) to direct attention externally can potentially keep the internal content of working memory unchanged; thereby the working memory content shall not be able to coordinate the selection of external stimuli. Moreover, researchers sought to verify the presence of a limited, shared attention resource, exogenous and endogenous, and its relation to working memory content. Conclusively, redraws of internally-oriented attention were hampered on account of time demands as well as cognitive load, followed by a lack in the refreshment of working memory contents and poorer working memory recognition performance. In total, high frequency external stimulus processing demands, limited both influence and retention of the internal content of working memory, thus explaining why working memory contents might capture visual attention in some situations but not others.

Sauce et al. [25] investigated on the types of attention and their relation to the learning abilities of mice. External attention has already been connected to selective or perceptual attention originating from external sources, whereas internal attention represents inhibition, self-control and reflective attention originating from internal sources. Exploratory factor analyses, followed by confirmatory factor analyses on the attentional patterns of the performance in four distinct tests of attention in mice, confirmed the binary mechanism of attentional systems. In addition, internal and external attentional systems can potentially explain some of the normal variation in the learning abilities of mice; Moreover, the efficacy of higher resistance against external sources of interference in mice was found to promote new learning.

3 Computational Models of Attentional Control Processes

Lu et al. [17] developed the noisy Perceptual Template Model (PTM) for predicting the effect of attention on visual tasks either through a signal enhancement mecha-

nism, distractor exclusion or through internal noise suppression. Any attention effect should manifest itself either in the low noise region, the high noise region or both. This model consists of four major components; a perceptual template, a multiplicative internal noise source, an additive internal noise source and a decision process. Depending on the task, the decision reflects either detection or discrimination. Furthermore, researchers studied attention mechanisms in a location-cued orientation discrimination task and the threshold signal contrast level for each subject was determined for all combinations of attention conditions and external noise levels. The results confirmed that location-cued attention alters discriminability of targets. More specifically, signal enhancement after the introduction of internal multiplicative noise affected individuals' performance in the task. The major benefit of the PTM model is the applicability of external noise manipulation as a means for testing the mechanisms of attention in a wide range of perceptual domains. Future research should further specify the nature of perceptual processing performed by the perceptual template and this addresses the need for manipulating the content of the external noise and its relation to internal noise.

Itti et al. [11] presented a bottom-up saliency model related to the “feature integration theory”. This framework can be used for rapid scene analysis through selecting a small number of interesting image locations to be analyzed by more complex and time-consuming object-recognition processes. Therefore, all feature maps feed in a purely bottom-up manner, into a master “saliency map”, which topographically codes for local conspicuity over the entire visual scene. However, only object features represented in an explicit way in at least one of the feature maps can lead to rapid detection independent of the number of distracting objects.

Itti et al. [9] have stressed five important computational trends with particular emphasis on bottom-up control of attentional deployment. The computation of saliency derives from low-level visual features in contrast to the contextual surround. Saliency is coded explicitly in cortex, thus many explicit saliency maps must be integrated to yield unitary behavior. Moreover, inhibition of return is a complex, object-based and dynamically, adaptive process, while eye movements need to be integrated into models. Last, attentional orientating and object identification relate to attentional deployment and top-down cues. Future research should centre on modeling of interactions between task demands and integrated top-down and bottom-up cues as well as on the interplay between attentional orienting and scene or object recognition.

Itti [8] presented an application of attention models to evaluate advertising designs and more specifically, to select the most adequate image for a cover of a magazine. Locations marked as highly salient by the model both attracted the gaze of a majority of potential customers and conveyed important information for the advertiser. Therefore, whichever candidate design had the highest average model-predicted salience over the important text messages, won.

McCarley et al. [18] developed a computational model of attention and situation awareness (A-SA) in an attempt to predict pilot errors during taxiway operations on the airport. The attention module was based on Bundesen's Theory of Visual Attention, presupposing that attention is allocated in varying quantities to objects and events within a scene. The conspicuity and information value of the item form the

attention allotment. Conspicuity refers to the physical perceptibility of the item, whereas information value encompasses the effect of the item on situation awareness. Thus, high situation awareness guides attention towards objects and events at the expense of items, degrading situation awareness. Therefore, expectancy on evidence seeking, that is an anticipated and sought by the pilot item, can consume attentional resources. Moreover, the attentional weight of an item declines according to an exponential decay function. As a result, the model can capture the degrading influence of the cognitive load on situation awareness. Furthermore, allocation of attention to new items during attentional processing of irrelevant stimuli shall lead to a faster decay rate for situation awareness as working memory processes are disrupted more than during the absence of attention demanding stimuli. Evidently, changes in situation awareness were manifested as changes in top-down attentional guidance, indexed by the amount of attention allotment between relevant and irrelevant items as well as in the navigational behavior, indexed by the probability of correct behavior at taxiway intersections.

Barnard et al. [1] presented an implemented model of the attentional blink effect relying on process exchanges between propositional meaning and a more abstract, implicational level of meaning; the latter represented and captured affect. Process algebra, a formal specification technique, was developed in order to build this computationally explicit model of human attention in cognitive-affective settings. Interacting Cognitive Subsystems (ICS) refer to a central engine of mentation, reflecting the distributed nature of executive control. In the ICS architecture, subsystems are themselves systematically decomposed into components with distinct internal functions, such as an array representing the input data to a subsystem, an image record and processes that transform data from one type of mental representation to another. With its explicit hierarchical structure and definition of processes that communicate within and between modules (buffered processing) in combination with its adherence to the distributed control hypothesis, ICS is ideal for realization through process algebra. Buffered processing is associated with focal awareness and the consequent allocation of limited attentional resources. Therefore, the buffer acts as a movable locus of control within distributed processing activity. The model runs and produces a blink effect that is the time duration for the buffer moving from one subsystem to another according to the features of the representations processed in real time, without requiring more detailed assumptions about perceptual, lexical or semantic representations.

Kasderidis et al. [12] created an autonomous robot control system based on an attention-processing model. In this framework, attention serves as the learning generator that controls the system in a multi-modal level as well as a global competition system among conflicting goals. The architecture has four processing levels; level one includes sensors and actuators, level two has pre-processing facilities, level three has schemata for atomic goals and finally, level four has schemata for composite goals. The proposed architecture was examined in terms of its effectiveness in detecting a danger in addition to learning effective motion models of route planning in a dynamic environment. The attention-based predictive agent enabled the agent to learn dynamically corrections to its poor actions as well as to automate the corrected behavior. It seems that attention controls the switching from the automated mode to learning

mode. Advanced control of the agent in a feature selection process could encompass a multi-modal system or an automatic creation of goals and response rules originating from the system, itself.

Bertel [2] proposes that anticipatory cognitive computing can integrate a live influx of behavioral data on attention recorded from a human reasoned into an artificial cognitive system. According to the author, a general cognitive architecture of attentional processing and its adjustment to inter-individual differences as well as eye tracking techniques and multi-modal data may serve future modeling of the production of overt and covert attentional shifts.

Reynolds et al. [24] created a normalization model of attention through embracing the capacity of the brain to increase sensitivity to faint stimuli presented alone and reduce the impact of task irrelevant distracters when multiple stimuli are presented. The three basic components of the model entail the stimulation field, the suppressive field and the attention field. Therefore, stimulus conditions and the subject's attentional strategy reshape the distribution of activity across the population of neurons, shifting the balance between excitation and suppression. In other words, the effect of attention should systematically shift from response gain to contrast gain by appropriate manipulations of the stimulus size and attention field size.

Neokleous et al. [20] designed and implemented a computational model of visual selective attention, the main control mechanism for keeping the brain system in stability. Selective attention filters out any irrelevant information while advancing any vital stimulation to higher cortical areas for further processing, leading to conscious perception. Top-down and bottom-up attentional processing addresses individuals' need to balance internal motivations and external stimulations. Endogenous attention is considered to be a volitional modulation of neural activity in the parietal and frontal lobes of the brain, corresponding to an object or a location in space and exogenous attention is automatic and relies on the sensory saliency of stimuli registered by sub cortical structures and the primary sensory cortices. Current computational model was built upon an appropriate system of dynamical equations, originating from the functionality of the brain and implementing spiking neural networks (SNN). The first stage simulates the bottom-up competitive neural interactions among visual stimuli, while the second stage entails modulations of neural activity based on semantics of the stimulus. Therefore, there is a constant interaction among the working memory, bottom-up and top-down attention depending on the rate and temporal coding of the incoming visual stimuli. The functionality of this model lies on the presence of a saliency map, analyzing an image through distinct characteristic maps and other specific operands as well as a winner-take-all neural network selecting the area of the image towards which attention is oriented. Moreover, endogenous attention is believed to manipulate the whole process through the synchronization of incoming stimuli with the goals of a specific task.

Borji et al. [3] made a review study on models of attention from psychological, neurobiological and computational perspective. Visual attention is confounded on the idea of a selection and a relevance mechanism, whether scene-driven bottom-up (BU) or expectation-driven top-down (TD), capable of working in real-time. The present review is centered on bottom-up computational models of saliency maps from any

image or video input. A saliency map is a topographic map representing conspicuousness of scene locations. The complexity in the distinction among the aforementioned models lies on the multi-sensory background of the human attention system; humans are capable of attending to multiple regions (between four and five) of interest simultaneously. Moreover, bottom-up and top-down attentions are combined to direct humans' attentional behavior. Therefore, attentional models that mimic humans' biological functions may well serve future research in predicting fixations over images and videos. A unified comparison framework in the evaluation metrics of attentional models, such as standard benchmark datasets and rigorous performance measures would enhance the whole field of attention modeling. More specifically, future models should improve the prediction accuracy of computational models through incorporating interactive, bottom-up and top-down factors. Thus, principled ways to model time varying task demands and covert and overt visual attention as well as building a map from fixations of other subjects over the same stimulus, surpass the scope of computer vision and enter the machine learning community.

Kaspar [13] investigated on the role of overt shifts in visual attention through recoding eye movements. More specifically, full-color computer-generated fractal images as well as urban and natural real-world scenes were implemented in order to look into both the bottom-up and top-down attentional processes influencing eye movements. Future research should center on motion saliency maps, saccade planning in combination with the weighting of scene features in a saliency map derived from top-down factors. In reference to top-down factors, such as arousal and valence, personality traits and motivational dispositions, time-dependent changes (when, how long and how strong) could produce interindividual differences on attentional control and thus, on eye movement behavior. Therefore, the interaction between stimulus-driven overt attention, emotional components and personality traits can explain much of the variance in eye movement behavior across and within subjects.

Omori et al. [21] proposed a symbolic thinking process forming the Q-learning procedure acquisition model using a computer simulation of a navigation problem. This mental procedure formation for problem solving is composed of a sequence of subfunctional module combinations based on a continuous computational process. More specifically, the combination process in human brains is realized by internal attention as a vector that selects necessary cortical areas and functional parts. Therefore, symbolic internal processes in computational circuits are structured according to a sequential program called "procedure". Thereby, in the "procedure" formation, the system traces the suitable functional parts as well as their combination at a specific task. The "procedure" becomes automatic when a suitable attention vector sequence is established, including finding and memorizing the activation vector for the functional parts of respective tasks. In the experiment, researchers used two types of agents; a basic agent and a developed agent, both designed to realize Q-learning and prediction-based learning. Moreover, as attention generator, a two-step sequence attention vector was implemented; the latter remained unchanged along the different tasks. As a result, self-organization in this model seems to be reciprocally related to various parameters of the learning system, such as an increase of available functional parts.

Poth et al. [23] interrelated monitoring for visual events to distinct components of attention. According to the computational Theory of Visual Attention (TVA), visual information processing is organized in separate processing stages. First, visual short-term memory (VSTM) maintains visual information temporarily and stores information about a respective object after the latter wins the competition among other possible categorizations of objects. In addition, two attentional mechanisms, a filtering and a pigeonholing mechanism, influence the processing speed of categorizations and thus enhance the probability of an object to be kept in visual short-term memory. Moreover, preparatory attentional processes monitor the environment for events to support event-based prospective memory. Therefore the present study examined the effects of monitoring for visual events on the distinct attentional components in a concurrent visual attention task as well as whether these effects varied along with the expected perceptual salience of the events. Visual processing speed, conscious perception, the top-down controlled selectivity and the laterality of attentional weighting were measured in order to examine expected interrelations under three conditions; the low- and high-salient event condition and the control condition. However, sustaining attention to a task as well as the visual attention component did not appear to be inter-related. Notwithstanding, by increasing the salience of events that need to be monitored, visual processing resources are redistributed in favor of an ongoing task and at the expense of an event-based prospective memory task.

Stollenga et al. [26] introduced Deep Attention Selective Networks (dasNet), modeling selective attention in deep Convolutional Neural Networks (CNNs) by allowing each layer to influence all other layers on successive passes over an image. The aforementioned Convolutional Neural Network offers special connections both bottom-up and top-down that modulate the activity of the convolutional filters, thus adapting the CNNs behavior. Therefore, the CNN through reinforcement learning and the Separable Natural Evolution Strategies (SNES), is able to direct selective internal attention to certain features extracted from images over multiple passes in order to improve the difficult, image-classification cases, not captured by the initially supervised training.

Mnih et al. [19] developed a novel recurrent neural network (RNN) model that can be trained using reinforcement learning methods to learn task-specific policies. In addition, it can control the amount of performing computation regardless of the input image size. It is formulated as a single recurrent neural network, which takes a glimpse window as its input and uses the internal state of the network to select the next location to focus on as well as to generate control signals in a dynamic environment. Therefore, the model considers attention-based processing of a visual scene as a control problem and is general enough to be applied to static images, videos or as a perceptual module of an agent that interacts with a dynamic visual environment. The current model is applicable to large scale object recognition and video classification.

Itti et al. [10] made a taxonomy of visual attention, encompassing the bottom-up, stimulus-driven, saliency-based attention and the endogenous, top-down, task-driven attention. The former operates in parallel throughout the entire visual field and assists in mediating pop-out interference, while the latter depends on the exact task at hand and on subjective visual experience as well as it takes longer to deploy and is voli-

tionally controlled. In addition, they typically involve cognitive reasoning aimed at rapidly extracting a task-dependent compact representation of the scene that can be used for further reasoning and planning of top-down shifts of attention and action. Most importantly, there are top-down models using fuzzy or probabilistic reasoning to explore how several sources of bottom-up and top-down information may combine.

4 Research Highlights

Visual attention resembles a spotlight, shedding light on individuals' external and internal stimuli simultaneously. Perceptual/external attention addresses the need for selecting and modulating sensory information, whereas central/reflective/internal attention sets priority on the manipulation of information in the working memory and the long-term memory. Therefore, behavior and decision making is the outcome of the reciprocal interaction between competitive stimuli and processes, either volitionally or not, sharing a common and limited cognitive resource.

More specifically, individuals' external attention processes, such as selection and modulation have an effect on working memory encoding, maintenance and recall. On the other hand, attentional control and working memory processes can bias external selection and modulation through a task-relevant search template for the guidance of visual search. In addition, external attention demanding tasks into shorter periods of time have an effect on one's ability to encode, maintain internally attended information and remove interfering representations. Therefore, working memory contents cannot coordinate the selection of external stimuli. Moreover, higher resistance against interference in mice promoted their new learning capability.

As a result, cognitive load on working memory due to emphasis on external attention cues has an effect both on knowledge encoding, structure of mental representations and working memory retrieval. Especially, knowledge encoding sets the background for efficient knowledge retrieval and manipulation. Thus, the internal attention mechanism has to be flexible in order to be able to perform knowledge encoding under stressful circumstances for individuals. By flexibility it is meant that individualized attentional strategies have to be enacted within modalities or domains in order to shift the balance between excitation and suppression. The question to be answered is how attentional patterns can be altered in favor of consciousness.

As a result, attentional control is fundamental for memory encoding, retrieval and manipulation in working memory. In addition, the original attentional foci leading to effective working memory function has a visuo-spatial print and it is stimuli-feature specific. The corresponding control processes for these foci on respective individuals could lead to a cross-level binding mechanism that may serve better understanding of human problem solving. Dynamic modeling of attentional shifts through analyzing a person's eye movements could reveal his spatial or diagrammatic reasoning and gradually, build on a dynamic model of focus in problem solving.

In any case, attention can be trained as a metacognitive and conscious process. Internal attention towards negative cognitions and strong emotions has already been found susceptible to mindfulness thinking exercises.

Standard benchmark datasets and rigorous performance measures can enhance the field of attention modeling. Incorporating interactive, bottom-up and top-down factors in computational models shall evoke the latter's prediction accuracy. Furthermore, anticipatory cognitive computing encompasses problem solving techniques, structural principles and cognitive processing mechanisms originating from human cognitive processing to generate cognitively-anticipatory behavior in an artificial cognitive system. After all, a cognitive model of what we expect to see is the basis for our percept.

5 Conclusion

Attention is the basis of human cognition and can be trained in favor of mental flexibility, learning and self-consciousness. More specifically, it has been found that working memory contents are reciprocally related to attentional control, shifting the balance between suppressing and prioritizing information processing. However, the features of the stimuli and individualized human characteristics, such as emotions, mood, and character can interfere with measuring attention. Therefore, research on eye movements and fixations could reveal attentional patterns revealing individualized attentional strategies regarding expectations and performance attributes. Thus, advanced psychometric data could be used for investigating the relation between attention and consciousness.

6 References

- [1] Barnard P. and Bowman H. (2003). Rendering information processing models of cognition and affect computationally explicit: distributed executive control and the deployment of attention. *Cognitive Science Quarterly*, 3(3): 297-328.
- [2] Bertel S. (2007). Some notes on the control of attention, its modeling and anticipatory cognitive computing. *American Association for Artificial Intelligence*.
- [3] Borji A. and Itti L. (2013). State-of-the-Art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(1): 185-207. <https://doi.org/10.1109/tpami.2012.89>
- [4] Chun M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia* 49: 1407-1409. <https://doi.org/10.1016/j.neuropsychologia.2011.01.029>
- [5] Chun M. M. Golomb J. D. and Turk-Browne N. B. (2011). A Taxonomy of External and Internal Attention. *Annu. Rev. Psychol.* 62: 73-101. <https://doi.org/10.1146/annurev.psych.093008.100427>
- [6] Gehring W. J. Bryck R. L. Jonides G. Albin R. L. and Badre D. (2003). The mind's eye, looking inward? In search of executive control in internal attention shifting. *Psychophysiology* 40: 572-585. <https://doi.org/10.1111/1469-8986.00059>
- [7] Hollingworth A. and Hwang S. (2013). The relationship between visual working memory and attention: retention of precise color information in the absence of effects on perceptual selection. *Phil Trans R Soc B* 368: 1-8. <https://doi.org/10.1098/rstb.2013.0061>
- [8] Itti L. Models of bottom-up attention and saliency. University of Southern California, Hedco Neuroscience Building HNB-30A, Los Angeles, CA 90089-2520.

- <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.9678&rep=rep1&type=pdf>.
- [9] Itti L. and Borji, A. (2015). Computational Models of Attention”, <https://arxiv.org/abs/1510.07182>.
- [10] Itti L. and Koch C. (2001). Computational modeling of visual attention. *Nature Reviews/Neuroscience*, 2: 194-203.
- [11] Itti L. Koch C. and Niebur E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11). <https://doi.org/10.1109/34.730558>
- [12] Kasderidis S. and Taylor J. G. (2003). Attention-based learning. *Attentional Agents and Robot Control*, KES Journal, IOS Press.
- [13] Kaspar K. (2013). What guides visual overt attention under natural conditions? Past and Future Research. *ISRN Neuroscience*, 1-8. Article ID 868491. <https://doi.org/10.1155/2013/868491>
- [14] Kerr C. E. Sacchet M. D. Lazar S. W. Moore C. I. and Jones S. R. (2013). Mindfulness starts with the body: somatosensory attention and top-down modulation of cortical alpha rhythms in mindfulness meditation. *Frontiers in Human Neuroscience*, 7(12): 1-15. <https://doi.org/10.3389/fnhum.2013.00012>
- [15] Kiyonaga A. and Egnér T. (2014). Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content. *Frontiers in Human Neuroscience* 8(670): 1-9. <https://doi.org/10.3389/fnhum.2014.00670>
- [16] Kiyonaga A. and Egnér T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychon Bull Rev.*, 20: 228-242. <https://doi.org/10.3758/s13423-012-0359-y>
- [17] Lu Z.-L. and Doshier B. A. (1998). External noise distinguishes attention mechanisms. *Vision Res.*, 38 (9): 1183-1198. [https://doi.org/10.1016/s0042-6989\(97\)00273-3](https://doi.org/10.1016/s0042-6989(97)00273-3)
- [18] McCarley J. S. Wickens C. D. Goh J. and Horrey W. J. (2002). A computational model of attention/situation awareness”, *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*. <https://doi.org/10.1177/154193120204601730>
- [19] Mnih V. Heess N. Graves A. and Kavukcuoglu K. (2014). Recurrent models of visual attention. <https://arxiv.org/abs/1406.6247>.
- [20] Neokleous K. C. and Schizas C. N. (2011). Computational Modeling of Visual Selective Attention. *Procedia Computer Science*, 7: 244-245. <https://doi.org/10.1016/j.procs.2011.09.030>
- [21] Omori T. and Ogawa A. Model of symbolic looking procedure acquisition process in navigation learning task, <http://www.lucs.lu.se/LUCS/085/Omori.pdf>.
- [22] Phillips I. (2012). Philosophical Perspectives. *Philosophy of Mind*, 26: 277-308.
- [23] Poth C. H. Petersen A. Bundesen C. and Schneider W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, 5(930): 1-11. <https://doi.org/10.3389/fpsyg.2014.00930>
- [24] Reynolds J. H. and Heeger D. J. (2009). The Normalization Model of Attention. *Neuron*. 61: 168-185.
- [25] Sauce B. Wass C. Smith A. Kwan S. and Matzel L. D. (2014). The external-internal loop of interference: Two types of attention and their influence on the learning abilities of mice. *Neurobiol Learn Mem.*, 116:181-192. <https://doi.org/10.1016/j.nlm.2014.10.005>
- [26] Stollenga M. Masci J. Gomez F. and Schmidhuber J. (2014). Deep networks with internal selective attention through feedback connections. *ICML 2014*: 1-13.
- [27] Tsuchiya N. and Van Boxtel J. (2013). Introduction to research topic: attention and consciousness in different senses. *Frontiers in Psychology*, 4(249): 1-4. <https://doi.org/10.3389/fpsyg.2013.00249>

- [28] Wu Dowd E. Kiyonaga A. Egner T. and Mitroff S. R. (2015). Attentional guidance by working memory differs by paradigm: An individual-differences approach. *Atten Percept Psychophys.*, 77: 704-712. <https://doi.org/10.3758/s13414-015-0847-z>

7 Authors

A. Drigas is with N.C.S.R. 'Demokritos', Institute of Informatics and Telecommunications, Telecoms Lab - Net Media Lab, Agia Paraskevi, 153 10, Athens, Greece (e-mail: dr@iit.demokritos.gr).

M. Karyotaki is with N.C.S.R. 'Demokritos', Institute of Informatics and Telecommunications, Telecoms Lab - Net Media Lab, Agia Paraskevi, 153 10, Athens, Greece (e-mail: karyotakimaria@gmail.com).

Article submitted 2019-01-21. Resubmitted 2019-03-17. Final acceptance 2019-03-27. Final version published as submitted by the authors.