

# Secondary Task Display Attributes – Optimizing Visualizations for Cognitive Task Suitability and Interference Avoidance

C. M. Chewar, D. Scott McCrickard, Ali Ndiwalana, Chris North, Jon Pryor, and David Tessendorf

Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

---

## Abstract

*We found that established display design guidelines for focal images cannot be extended to images displayed as a secondary task in a dual-task situation. This paper describes an experiment that determines a new ordering guideline for secondary task image attributes according to human cognitive ability to extract information. The imperative for alternate guidelines is based on the difference in an image's ability to convey meaning, which decreases when moved from a focal to a secondary task situation. Secondary task attribute ordering varies with the level of degradation in the primary task. Furthermore, attribute effectiveness may be particular to types of visual operations relating to cognitive tasks.*

Keywords: dual-task, visualization, information design, image attribute

---

## 1. Introduction

As humans increasingly integrate computer systems into their daily tasks, and as desktop systems expand in processing capabilities, there seems to be a universal desire to use these technologies to optimize delivery of information. Having the right information at the right time is certainly helpful—often vital—for any decision making process. However, we recognize that too much information, or information that is presented in a way that overwhelms the human sensory or cognitive system can also fail to inform, impairing a decision making process. Therefore, design of computer system interfaces must properly blend human cognitive and perceptual capabilities and preferences within system constraints.

This goal is central to studies within the realm of human-computer interaction, integrating many other disciplines for the express purpose of improving interface and visualization capabilities of computing systems. Since the necessity to recompense human limitations drives this area of research, study through a human factors lens can be helpful. Some aspects of experimental psychology linking human cognition and perception are useful to refine understanding of these particular human factors. Since

properly designing systems that address findings of human factors research is generally the application of human factors engineering, or more narrowly engineering psychology, it is also useful to synthesize and apply established ideas from that domain. This cross-disciplinary approach represents a pursuit for visualization design guidance that allows programmers to create computer interfaces that proffer insight rather than impose information glut.

A specific research area of growing importance within these efforts is the study of dual-task interface and visualization design. In many ways, desktop computer users are not content or able to simply devote their full attention to a single application or task. Instead, they need to monitor and maintain awareness of information through joint or secondary tasks. Simple examples of dual-task situations include editing a document while monitoring email programs, news delivery agents, instant messengers, and stock tickers. However, critical primary tasks—like driving vehicles, performing surgery, or even engaging in law enforcement or combat—can also be enhanced with vehicular displays, surgery support monitors and military situational awareness systems. The challenge to this research area is to balance effective information

presentation with introduction of unwanted primary task distraction. Since a secondary task often must conform to requirements seeking minimized distraction, even the most accepted focal task design principles merit consideration for sustained applicability in this altered condition. One part of these many principles includes the guidelines for effective use of various attribute-encoding schemes to communicate information for the fulfillment of different cognitive tasks. These principles must be reestablished or identified for secondary task display interfaces and visualizations expected to create actionable insight while minimizing distraction.

This study empirically investigates a dual-task system, working toward such guideline reestablishment or identification. The initial question is whether a user's ability to perform an information extraction task using a visualization on a desktop computer with certain attribute encoding is different when the task is a single task in the user's focus compared to a secondary task in a dual-task situation. If there are no performance differences, to include introduction of distraction to the primary task in the dual-task situation, then this indicates that focal guidelines for these attribute encodings are extensible to such a dual-task condition. However, if differences exist, then new guidelines for attribute use must be thoroughly investigated and enunciated.

## 2. Related Work

The basis that this work investigates for dual-task applicability comes from Cleveland and McGill's conclusions that provide an accepted guideline for the presentation of visual data in quantitative tasks, based on psychophysical theory and experimentation. They recognize visual data as elementary perceptual tasks, described as graph attributes, some of which convey information better than others. Attribute effectiveness guidelines facilitate design of display interfaces and visualizations that effectively communicate information and create insight. The ordering of graph attributes which Cleveland and McGill empirically provide is: position along a common scale, position along nonaligned scale, length, direction, angle, area, volume, curvature, shading, and color saturation [3]. Mackinlay extends this list to capture nonquantitative data [5], resulting in the inclusion of more attributes and orderings for nominal and ordinal data. For both nonquantitative perceptual tasks, color is better than area, although position remains the best. These graphical design principles are established for visualizations that are the object of a user's focal attention.

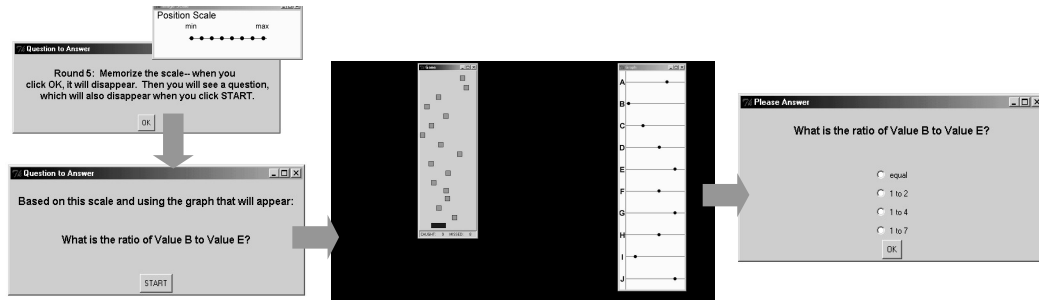
Cleveland extends his thinking about attribute effectiveness, relating specific cognitive task requirements and visual operations. He recognizes three types of pattern perception operations that form all operations of physical information extraction from graphics: symbol *detection*,

*assembly* (or grouping), and *estimation* (discrimination, ranking, and ratioing) [2]. Several general assertions entwine graph attributes and perception operations. Detection tasks are best supported by single curves or line segments; filled circles that may overlap hinder detection tasks. He states that color can be used to establish categories that enhance assembly, since assembly is enhanced by symbols that have strong boundaries (non-overlapping area). Using color can also provide quantitative encoding that increases estimation efficiency. Position in relation to a reference grid and dot plots with ordered categories improves all pattern perception operations, especially estimation. He criticizes circular area encoding in general, since it fails to "provide efficient detection of geometric objects that convey information about differences of values."

Wickens et al. introduce discussion of fundamental cognitive processes as well—*search* and *compare*—"that may be supported or inhibited by specific graphical renderings" [12]. This notion of search tasks seems to include Cleveland's detection, assembly, and estimation-discrimination operations. Compare tasks roughly equate to estimation-ranking and estimation-ratioing. Wickens and Hollands examine relative attribute effectiveness as a function of human ability to conduct a parallel search among color variation, as opposed to more timely decoding of other attribute encodings searched in serial [11]. Lohse takes a similar approach, stating that since color is detected and organized in parallel during pre-attentive visual processing, it is a more efficient encoding than area, since shape is detected serially [4].

Important research on elements of dual-task display effectiveness has emerged in recent years, although nothing appears to apply directly to an ordering of visual attributes. Mori and Hayashi's work establishes peripheral task causes of primary task interference in multi-window systems [8]. Wickens and Hollands discuss primary task performance degradation in dual-task situations in terms of resource allocation to secondary tasks and adaptation consequences for excessive workload [Wickens 2000]. Rock and Mack also examine divided attention with respect to parallel and serial pre-attentive processing [9]. Others investigate properties of secondary tasks. McCrickard et al. find that effectiveness of different types, sizes and speeds of secondary task text displays relate to different levels of performance expectations—either identification, or higher level comprehension and memorability [7]. Maglio and Campbell conclude that constantly scrolling text should be minimized, since scrolling distracts more than text that discretely appears and disappears [6]. Bartram shows effective uses of motion in displays, particularly with respect to information presence signaling, information search and association, and filtering or linking of spatially distributed objects [1].

The following experiment builds on all of the work discussed above, seeking understanding of dual-task



**Figure 1:** Screen shots of experimental platform. A round begins with scale display; scale disappears and question is displayed; question disappears and game begins—alone for eight seconds (not shown); eight seconds of game playing (left side of screen) and image display (right side of screen); eight final seconds of game playing only (not shown); answer input.

visualization guidelines. However, three major works are the foundation of this study: Cleveland and McGill's efforts at attribute ordering [3] and the cognitive task conceptualizations from Cleveland [2] and Wickens et al. [12].

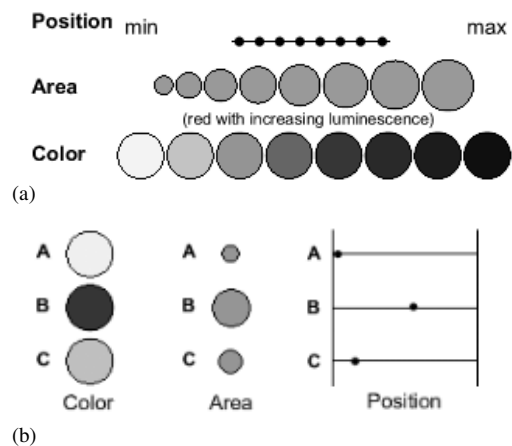
### 3. Experimental Design

In order to empirically test relative attribute effectiveness, a participant plays a simple, yet demanding game on a desktop computer. Scripted, timed events present experimental conditions and record participant performance throughout the experiment. During the game playing, which occurs on the left portion of the screen, a single image with similar dimensions and brightness as the game appears for eight seconds on the right screen edge (see Figure 1). The eight-second display time allows data within attended and ignored locations to be reliably and accurately detected [10].

The game playing continues while the participant scans the image for information—the answer to a question asked before the round begins. Each instance of the experiment includes eighteen rounds—nine dual task rounds (game-playing and image viewing) as well as nine focal (game-free) rounds. Both treatments require viewing images and answering questions. The participants are 93 undergraduate computer science students, who received class credit.

Six versions of the program implement a Latin square experimental setup testing the independent variables (three attributes, two conditions—single (focal) and dual task). Three base versions differ only in attribute presentation order. Each of these three versions provides two test iterations—one that starts with the dual task and finishes with the focal images, and the other that reverses this

sequence. Figure 2 shows attribute scales and encoding schemes.



**Figure 2:** Attribute scales and encoding schemes used in the experiment are shown (a). Relative increases within attribute values are uniform. Participants are shown the applicable scale before the start of each round. Each of the three attribute images (b) encodes the same dataset values {1, 5, 2}. In a given round, one of these images is presented as a secondary task. Only three values are shown here, but the experiment's images encode ten.

Game rounds cycle through three different question types (identification of displayed minimum/ maximum values, ratios, or comparison counts) for a single graphically encoded dataset. Regardless of version and attribute encoding, round questions and answers appear in constant order. For instance, images like those in Figure 2 are used

in all six versions as the first graph type, but the question (“what is the minimum value?”) and answer (A) do not vary. However, two versions (one version displaying the graph as a focal task, the other as a secondary task) encode this first dataset with position, two others use color, and the final two use area. After testing a single graph with three questions, the dataset and encoding scheme change.

#### 4. Results

The initial research objective for this experiment seeks to evaluate focal guideline applicability to dual-task conditions for effective use of position, color, and area attribute-encoding schemes. Guidelines must optimize probability of information communication through the fulfillment of various cognitive tasks. Improper guidelines can result in ineffectively designed visualizations and display interfaces. Rather than allowing information monitoring and awareness, which could lend enormous efficiency gains, poor multi-task displays can inhibit, rather than enhance, primary task performance. Results discussed below show focal guidelines do not hold for secondary tasks within a dual-task setup. Furthermore, a user’s question-answering performance drops significantly when an information extraction task is a secondary task rather than a single task in his focus. Both findings indicate proper selection of these three attributes for dual-task system design cannot rely on established guidelines governing focal visualization. Therefore, the majority of results in this section relate to the establishment of new guidelines. Our analytical approach includes evaluation of attribute effectiveness according to answer correctness by question type at various levels of primary task performance. Since each of the six experiment version samples were not significantly different than the population mean for any measured aspect of test performance, we consider the entire population for applicable results.

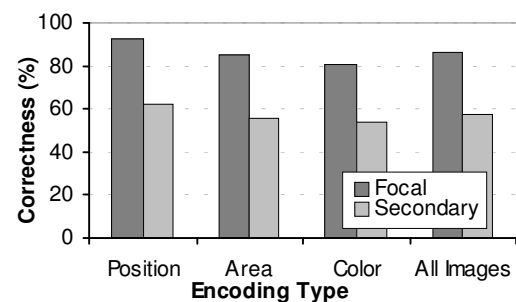
##### 4.1. Evaluation of Focal Guidelines

As we begin evaluation of focal guideline applicability to dual-task conditions, we expect part of the data should echo the ordering that Cleveland and McGill set forth [3]. Specifically, participants’ correctness of answers from the game-free conditions (which only provide the scale and question, the encoded dataset image, and the opportunity to input an answer) should correspond to an ordering of position, area, and then color. This is actually the case ( $F(2,277)=7.91$ ,  $MSE=.409$ ,  $p=.00046$ ), indicating that our attribute encoding scheme and experimental setup could at least replicate the part of Cleveland and McGill’s results under investigation.

The next step involves a comparison of correctness based on insight gained from focal images to correctness

facilitated by secondary task images. In the dual-task condition, answer correctness requires maintenance of a certain level of primary task (game) performance while the image is displayed. Evaluating a secondary task in a dual-task system with this method, performance effect on both tasks is considered. The mere concept of a dual-task situation mandates the evaluation of both tasks—if one or the other were not important, this removes motivation to optimize dual-task design; the system may as well be considered a single-task (focal) situation. However, dual-task systems must minimize negative effects created by the causal relationship between two tasks. Primary task *degradation* expresses change in game performance during image display period in relation to average game performance before and after the image appears. Answer correctness (secondary task measure) is evaluated for any round meeting a given primary task degradation threshold (*acceptable degradation*).

Figure 3 shows data comparing correctness levels, allowing acceptable degradation through 25 percent. However, for all degradation levels between zero and fifty-one percent, there are significant differences in the two results ( $z$ -scores range from 13.189 to 1.965,  $n=93$ ).



**Figure 3:** Resulting answer correctness percentages for images displayed as a focal task (displayed without any other objects on the screen) are compared to results from secondary task images (displayed while user continues playing a game). In this case, participants maintained 75 percent of game performance (acceptable degradation was 25 percent) during image display period for answer correctness to be evaluated.

In other words, given two identically encoded images—one in a user’s focus and the other displayed as a secondary task—a user is unable to extract information from the secondary display as effectively and/or without distracting their ability to adequately maintain primary task performance. Similarly, the two activities (extracting focal image information and extracting secondary task image information) are different. The implication of this result is that there is no reason to expect focal guidelines to hold for

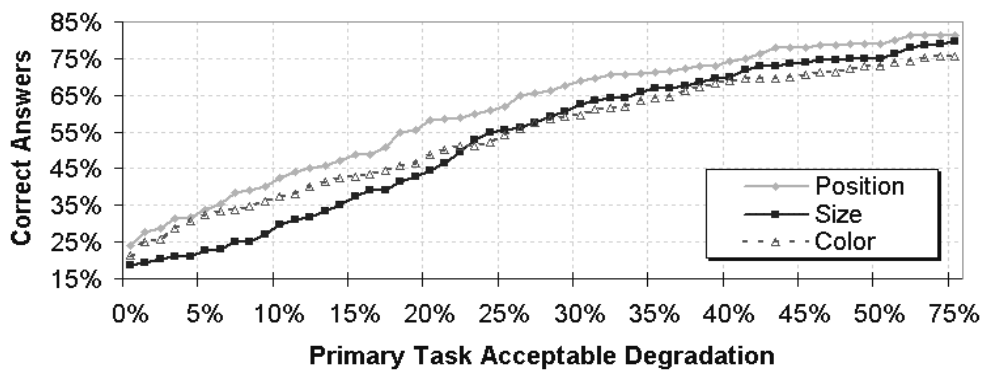
secondary tasks within a dual-task set-up. The following section shows that this implication is true, and that the attribute effectiveness ordering varies with levels of acceptable degradation in the primary task.

**4.2. Investigation and Establishment of Dual-Task Guidelines**

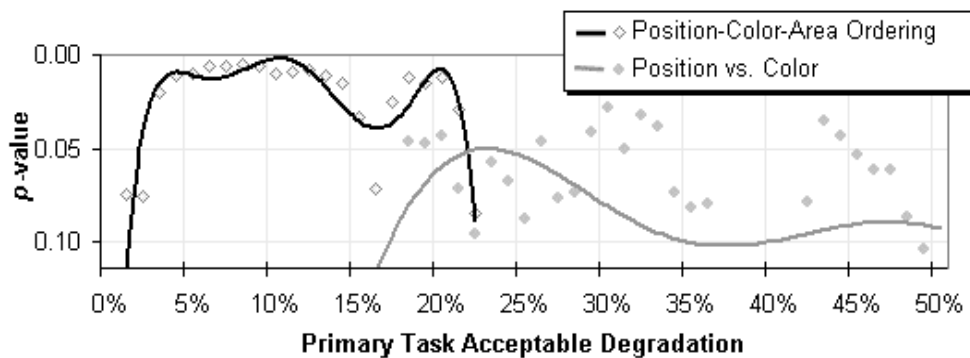
To determine attribute effectiveness ordering in the dual-task condition, we use the dual-task scoring system described above: answer correctness is evaluated for any round within a given level of primary task degradation. This evaluation allows identification of correctness averages by attribute type through any desired degradation level (see figure 4).

Given these correctness averages, between-group confidence levels are established with multivariable analysis of variance (ANOVA) tests. Significant results are depicted as points in Figure 5. Regressed confidence trendlines (sixth order) also in Figure 5, show a strong pattern of communication difference between secondary task images encoded with different attributes.

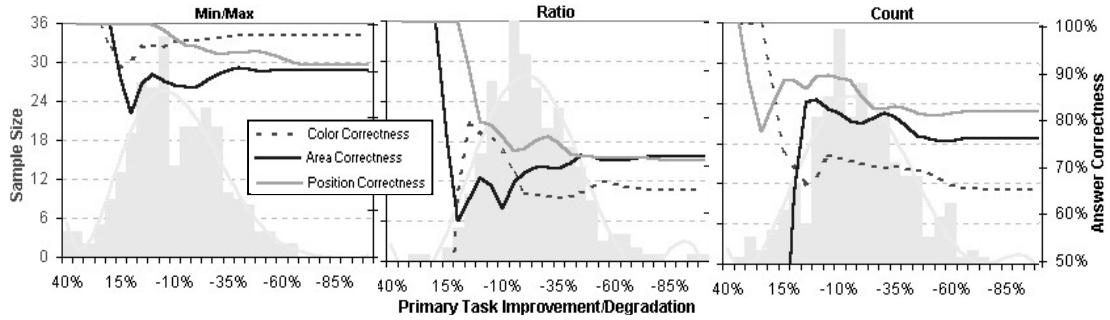
Furthermore, this ordering pattern varies in effectiveness with acceptable degradation and can be completely ordered at low degradation levels—between three and twenty-two percent—position, color, and then area. As degradation increases past 23 percent, certainty is lost about the distinction between position and area, but position is still a better facilitator of understanding than color. Supporting confidence levels for this result come from two-tailed z-scores from difference comparisons of sample mean



**Figure 4:** Answer correctness by attribute type for levels of acceptable degradation. Note that levels of acceptable degradation cumulate from the left side of the figure (i.e., subjects that meet five percent acceptable degradation include those at zero through four percent, and five).



**Figure 5:** Solid trendlines show that levels of significance for attribute ordering vary with degree of acceptable primary task degradation.



**Figure 6:** Each chart shows information about a single question type. Left y-axis and histogram data show sample size distribution at various levels of primary task performance changes. Right y-axis and trendlines (2-period moving averages) show variation of answer correctness according to primary task performance changes.

position and color performance.

### 4.3. Dual-Task Guidelines According to Cognitive Task

Although this guideline is useful by itself for design of secondary task images, our experiment data also lends itself well for analysis of attribute effectiveness for different cognitive tasks, such as those discussed by Cleveland [2] or Wickens et al. [12], summarized in the Related Work section above. The results that follow apply to attribute decoding success observed by question type correctness.

Our experiment uses three different question types to test information communication from the images (see Figure 2 for an example image). Each attribute is tested with each question, under both focal and secondary task conditions. Therefore, we can find correctness averages according to attribute type and question type, while filtering dual-task performance based on acceptable degradation. Figure 6 shows a summary of this data.

Clearly, most participants perform best on min/max questions, then count questions, and then ratio questions.

Also, count and ratio questions appear to create higher degradation levels in the primary task than min/max questions create. This is consistent with Cleveland’s classification—detection tasks should be faster and more accurate than estimation tasks.

Table 1 summarizes zones where significant attribute orderings occur for each question type/cognitive task. “Low Degradation” includes levels of primary task degradation less than seventeen percent, while “High Degradation” includes levels greater than seventeen percent. Note that attribute orderings at low degradation levels are the same for min/max and ratio questions, while orderings at high degradation levels are the same for ratio and all count questions. This possibly indicates similarities between two cognitive tasks, when performed at different levels of attention. For example, if a participant maintains high game performance (and answers the question correctly), we can infer that the secondary task did not require much attention to properly decode.

The position—color—area ordering apparent for the detection tasks and the estimation-ratioing tasks at low degradation matches Mackinlay’s attribute ordering for nominal and ordinal perceptual tasks [5].

Cognitive Task	Low Degradation	High Degradation
Find Min/Max (detection)	position—color—area	none
Determine Ratio (estimation- ratioing)	position—color—area	position—area—color
Count Comparison Matches (estimation-compare)	position—area—color	position—area—color

**Table 1:** Significant Attribute Ordering, by Cognitive Tasks ( $p < .05$ )

## 5. Conclusions and Further Work

Several conclusions are evident from these results. First, users are unable to extract information from the secondary display as effectively and/or without distracting their ability to adequately maintain primary task performance. Although display design for focal tasks is known to be maximized with use of position, area, and then color, these guidelines are not necessarily applicable to dual-task conditions. Since attribute effectiveness in secondary tasks varies with the amount of degradation introduced (allowed) on the primary task, display attributes should be selected after specification of acceptable amounts of primary task performance degradation. Designers of vehicular displays and other systems supporting a critical primary task would certainly want to consider attribute ordering at the lowest levels of acceptable primary task degradation. Therefore, selection should be guided by relevant attribute ordering in Figure 4, or Table 1 if cognitive tasks are known. One general conclusion is asserted—considering the superiority of position in all focal and dual-task orderings, information should be conveyed in terms of relative position whenever possible to allow optimal probability for accurate communication and primary task sustainment.

Further work should include evaluation of other display attributes and combinations of attributes, to include Cleveland and McGill's complete list. More narrow research on effectiveness of different position encodings may also be useful, especially in light of the conclusion stated above. However, the human ability to pre-attentively process color, which allows decoding efficiency, is apparent in these results and may have exciting implications for minimizing dual-task distraction. Since this particular experiment limited color encoding to incremental instances of red luminescence, it may also be useful to investigate relative effectiveness of other hues. Categorizing ideal display properties with cognitive tasks is also an area of immense research potential. Synthesis of these types of design guidelines allows programmers to create interfaces best suited for the human user and the required task.

## Acknowledgments

Special thanks goes to Jacob Somervell for his advice and experimental platform.

## References

1. L. R. Bartram, "Enhancing Information Visualization with Motion", Unpublished dissertation. School of Computer Science, Simon Fraser University (2001).
2. W. S. Cleveland, *The Elements of Graphing Data*. Hobart Press, Summit, New Jersey, 1994.
3. W. S. Cleveland, R. McGill, "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods", *Journal of American Statistical Association* **79**(387): 531-554, Sept 1984.
4. J. Lohse, "A Cognitive Model for the Perception and Understanding of Graphs", In: *Human Factors in Computing Systems—Reaching Through Technology (Proc. of CHI '91)*, 137-144, 1991.
5. J. Mackinlay, "Automating the Design of Graphical Presentations of Relational Information", *ACM Transactions on Graphics* **5**(2): 110-141, April 1986.
6. P. P. Maglio, C. S. Campbell, "Tradeoffs in Displaying Peripheral Information", In: *Human Factors in Computing Systems (Proc. of CHI 2000)*, 241-248, 2000.
7. D. S. McCrickard, R. Catrambone, J. T. Stasko, "Evaluating Animation in the Periphery as a Mechanism for Maintaining Awareness", In: *IFIP TC.13 Conference on Human-Computer Interaction (Proc. of INTERACT 2001)*, 148-156, July 2001.
8. H. Mori, Y. Hayashi, "Visual Interference With User's Tasks on Multi-window Systems", In: *Proceedings of the Fifth International Conference on Human-Computer Interaction*, 2:80-85, 1993.
9. I. Rock, A. Mack, "Attention and Perceptual Organization", In: Ballesteros, S. (ed.): *Cognitive Approaches to Human Perception*. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1994.
10. T. Weber, A. Kramer, O. Karni, "Tracking Visual Attention With Event-related Brain Potentials" In: *ARL Federated Laboratory Advanced Displays and Interactive Displays Consortium, Advanced Displays and Interactive Displays Second Annual Symposium*, 33-38, February, 1998.
11. C. D. Wickens, J. G. Hollands, *Engineering Psychology and Human Performance*. 3<sup>rd</sup> edn. Prentice Hall, Upper Saddle River, New Jersey, 2000.
12. C. D. Wickens, D. H. Merwin, E. L. Lin, "Implications of Graphics Enhancements for the Visualization of Scientific Data: Dimensional Integrality, Stereopsis, Motion and Mesh" In: *Human Factors*, **36**:44-61, 1994.