Copy No. _____

Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

# Evaluation of New Visualization Approaches for Representing Uncertainty in the Recognized Maritime Picture

*Michael Matthews, Lisa Rehak, Julie Famewo, Tamsen Taylor and Jeremy Robson*
*Humansystems® Incorporated*

*Humansystems® Incorporated*
*111 Farquhar St., 2nd floor*
*Guelph, Ontario  N1H 3N4*

*Contract Project Manager:  Michael Matthews,  519-836-5911*

*Contract Number:  W7707-054996/002/QCL*

*Call-up Number:   W7707-078067*

*Contract Scientific Authority:  Anna-Liesa S. Lapinski,  902-426-3100  ext 180*



## Defence R&D Canada – Atlantic

Contract Report
DRDC Atlantic CR 2008-177
October 2008

Canada

This page intentionally left blank.

# Evaluation of New Visualization Approaches for Representing Uncertainty in the Recognized Maritime Picture

Michael Matthews
Lisa Rehak
Julie Famewo
Tamsen Taylor
Jeremy Robson

Prepared By:
Human*systems*® Incorporated
111 Farquhar St., 2nd floor
Guelph, ON N1H 3N4

Contract Project Manager: Michael Matthews, 519-836 5911
Contract No: W7701-054996/002/QCL
Call-up No: W7707-078067
Contract Scientific Authority: Anna-Liesa S. Lapinski, 902-426 3100 ext 180



## Defence R&D Canada – Atlantic

Contract Report
DRDC Atlantic CR 2008-177
October 2008

Principal Author

Michael Matthews, PhD

Humansystems® Incorporated

Approved by

*Original signed by Mark McIntyre for*

James S. Kennedy

Head, Maritime Information and Combat Systems Section

Approved for release by

*Original signed by Ron Kuwahara for*

James L. Kennedy

Chair, Document Review and Library Committee

This report and the research described therein were completed under the 11h Maritime Domain Awareness Thrust, under project 11he, Information Visualization and Management for Enhanced Domain Awareness in Maritime Security. Standing offer contract No. W7701-054996/002/QCL supports the Joint Command Decision Support for the 21st Century (JCDS 21) TD.

In conducting the research described in this report, the investigators adhered to the policies and procedures set out in the Tri-Council Policy Statement: Ethical conduct for research involving humans, National Council on Ethics in Human Research, Ottawa, 1998 as issued jointly by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

# Abstract

This report documents the literature review and experimentation used to develop and assess visualization options to represent uncertainty in the Recognized Maritime Picture (RMP), which is the visual representation of the surface vessel picture for the Canadian maritime Area of Interest (AOI). Specifically, visualization options for the uncertainty with regards to the identity, spatial position and time lateness of surface contacts and the quality and time lateness of the sensor coverage were developed and assessed using computer-based experiments at the Human*systems* (HSI®) Test Lab. Two icons (Rectangle design and “Lego” design) were developed to display uncertainty related to the surface contacts, in addition to background swaths with two features (fill and border) to display sensor coverage uncertainty. Search times and accuracy were explored through 6 experimentation sessions with 11 participants. The results showed a small search time advantage for the Rectangle design and small performance differences among the different designs for sensor coverage. Participants rated the workload associated with using the designs as low. All of the design options evaluated are considered to be suitable candidates for future evaluation by the operational community. This work was conducted as part of the Information Visualization and Management for Enhanced Domain Awareness in Maritime Security Applied Research Project within the Defence Research and Development Canada (DRDC) Maritime Domain Awareness (MDA) research thrust.

# Résumé

Ce rapport rend compte de l’analyse documentaire et de l’expérimentation qui ont servi à élaborer et à évaluer des moyens visuels de représenter l’incertitude de la situation maritime générale (RMP), c’est-à-dire la représentation visuelle des navires de surface qui empruntent la zone d’intérêt maritime du Canada. Plus précisément, on a élaboré et évalué des moyens de représenter visuellement l’incertitude en ce qui concerne l’identité, l’emplacement et la tardivité des contacts de surface et la qualité et la tardivité de la couverture des capteurs, dans le cadre d’expériences informatisées réalisées au laboratoire d’essai de Human*systems* (HSI®). On a conçu deux icônes (le modèle Rectangle et le modèle Lego) pour illustrer l’incertitude relative aux contacts de surface, en plus de champs de fond (bandes) comportant deux attributs (remplissage et bordure) permettant d’indiquer l’incertitude de la couverture des capteurs. Le temps de recherche et l’exactitude des données ont été examinés lors de six séances d’expérimentation qui ont réuni 11 participants. Les résultats montrent que le temps de recherche est légèrement réduit lorsqu’on utilise le modèle Rectangle et que le rendement des différents modèles applicables à la couverture des capteurs est variable. Les participants ont évalué comme faible la charge de travail associée à l’utilisation des modèles. Ils estiment que tous les modèles conceptuels examinés méritent d’être évalués plus en profondeur par des utilisateurs. Ces travaux ont été réalisés dans le contexte du projet de recherche appliquée intitulé Visualisation et gestion de l’information pour accroître la vigilance dans le secteur maritime, qui relève du vecteur de Recherche sur la connaissance du secteur maritime (CSM) de Recherche et développement pour la défense Canada (RDDC).

This page intentionally left blank.

# Executive summary

## Evaluation of New Visualization Approaches for Representing Uncertainty in the Recognized Maritime Picture:

**Michael Matthews; Lisa Rehak; Julie Famewo; Tamsen Taylor; Jeremy Robson; DRDC Atlantic CR 2008-177; Defence R&D Canada – Atlantic; October 2008.**

**Introduction**

This report documents the literature review and experimentation used to develop and assess visualization options to represent uncertainty of contacts and sensor coverage in the Recognized Maritime Picture (RMP) which is the visual representation of surface vessels in the Canadian maritime Area of Interest (AOI). As part of the Defence Research and Development Canada (DRDC) Maritime Domain Awareness (MDA) research thrust, Information Visualization and Management for Enhanced Domain Awareness in Maritime Security, this project was aimed at identifying visualization options to ease RMP operator's assessment of the information about surface contacts (tracks) within the RMP.

Specifically, the objective was aimed at visualization techniques to represent the uncertainty inherent in the information presented in the RMP. In particular, information to be represented included uncertainty due to time, positional uncertainty and identity uncertainty of the tracks, and individual coverage uncertainty of the sensors and information sources (quality and time-lateness).

**Results**

Discussions between DRDC[1] and HSI®, in addition to the review of the literature, led to the development and selection of two icons (referred to as the Rectangle and Lego designs) to represent uncertainty related to the surface contacts, in addition to background patches (swaths) with two features (fill and border) to display sensor coverage uncertainty.

Analyses indicated faster search times for the Rectangle icon compared to the Lego icon. Search times were also influenced by the swath, with icons being selected slightly faster in grey shaded swaths (light or dark grey), rather than hashed swaths (fine or course grid). Contacts were selected faster when the swath border, which indicated time lateness, was in line format (broken or solid lines) compared to colour format (green or gold border).

Accuracy was high for all visualization techniques, with no significant differences between the different designs of icons and background swaths.

**Significance**

This project is the first to develop symbology that could be realistically integrated into a future display for the RMP to represent contact and sensor uncertainty. Such symbology would allow users of the RMP to (i) quickly locate surface contacts of interest for which the underlying metadata may be uncertain, or (ii) determine the reliability of sensor coverage for a particular

[1] References to DRDC concern the PA, Liesa Lapinski and co-investigator Sharon McFadden of DRDC Toronto.

area. The display symbology assessed through this project shows significant promise based on the ease with which it can be learned, searched and identified, suggesting that it should be further evaluated by the RMP operational community.

**Future Plans**

The next step in the project will be to take the prototype design concepts, marry them to real RMP (unclassified) data and develop an interactive display that closely resembles what operators now see for the RMP. This display will be taken to the Joint Regional Operations Centers for evaluation by operators and management, with a view to collecting quantifiable data that will reflect the suitability of the prototypes to support actual operational tasks.

# Sommaire

## Évaluation de nouvelles techniques de visualisation permettant de représenter l'incertitude dans la situation maritime générale

**Michael Mathews; Lisa Rehak; Julie Famewo; Tamsen Taylor; Jeremy Robson; RDDC Atlantique CR 2008-177; R & D pour la défense Canada – Atlantique; octobre 2008**

**Introduction**

Ce rapport rend compte de l'analyse documentaire et de l'expérimentation qui ont servi à élaborer et à évaluer des moyens visuels de représenter l'incertitude des contacts et de la couverture des capteurs dans la situation maritime générale (RMP), c'est-à-dire la représentation visuelle des navires de surface qui empruntent la zone d'intérêt maritime du Canada. Dans le contexte du projet intitulé Visualisation et gestion de l'information pour accroître la vigilance dans le secteur maritime, qui relève du vecteur de Recherche sur la connaissance du secteur maritime (CSM) de Recherche et développement pour la défense Canada (RDDC), on a cherché à déterminer quelles méthodes de visualisation permettraient de simplifier l'évaluation par l'utilisateur de l'information concernant les contacts de surface (objets de poursuite) obtenue de la RMP.

Les travaux étaient axés sur les techniques de visualisation permettant de représenter l'incertitude inhérente à l'information affichée dans la RMP. Les données à représenter étaient entre autres l'incertitude attribuable à des facteurs de temps, l'incertitude attribuable à la position et à l'identité des objets de poursuite, et l'incertitude de la couverture de chacun des capteurs ainsi que des sources d'information (qualité et tardivité).

**Résultats**

Les échanges entre RDDC[2] et HSI®, en plus de l'analyse documentaire, ont abouti à l'élaboration et à la sélection de deux icônes (appelées respectivement les modèles Rectangle et Lego) servant à représenter l'incertitude relative aux contacts de surface, ainsi que de champs de fond (bandes) comportant deux attributs (remplissage et bordure) permettant d'illustrer l'incertitude de la couverture des capteurs.

Les analyses montrent que le processus de recherche est plus rapide à l'aide de l'icône Rectangle que de l'icône Lego. Le temps de recherche est également influencé par la bande; la sélection des icônes était un peu plus rapide au moyen des bandes teintées en gris (gris pâle ou gris foncé) qu'à l'aide des bandes hachurées (grille à mailles fines ou larges). Les contacts étaient sélectionnés plus rapidement lorsque la bordure de la bande, qui indiquait les retards dans le temps, était affichée sous forme de ligne (ligne discontinue ou pleine) que par une couleur (bordure verte ou or).

Le taux d'exactitude était élevé pour toutes les techniques de visualisation. On n'a relevé aucune variation significative entre les différents modèles d'icônes et de champs de fond.

---

[2] RDDC désigne ici la chargée de projet, Liesa Lapinski, et la collaboratrice Sharon McFadden de RDDC Toronto.

**Pertinence**

Ce projet est le premier à élaborer une symbologie que l'on pourrait réellement envisager d'intégrer à un affichage éventuel de la RMP pour illustrer l'incertitude par rapport aux contacts et aux capteurs. Une telle symbologie permettrait aux utilisateurs de la RMP de (i) repérer rapidement des contacts de surface qui présentent un intérêt et pour lesquels les métadonnées sous-jacentes pourraient être incertaines, ou (ii) déterminer la fiabilité de la couverture des capteurs dans un secteur donné. Les symboles d'affichage évalués dans le cadre de ce projet sont fort prometteurs étant donné la facilité avec laquelle on peut les apprendre, les chercher et les identifier, ce qui implique que les utilisateurs de la RMP en milieu opérationnel auraient intérêt à les évaluer plus en profondeur.

**L'avenir**

La prochaine étape du projet sera d'assortir les concepts prototypes à des données réelles (sans classification) concernant la situation maritime et d'élaborer un affichage interactif très semblable à la RMP que les utilisateurs consultent en ce moment. Cet affichage sera installé dans les centres d'opérations régionales interarmées et soumis à l'évaluation d'utilisateurs et de gestionnaires, dans l'optique de recueillir des données quantifiables qui témoigneront de l'utilité des prototypes à l'appui de tâches opérationnelles réelles.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

We would like to thank Lt(N) (ret) J. R. Rafuse for providing valuable information that contributed to the development of design concepts for this project and for detailed background information on how operator tasks are affected by uncertainty in the contact data.

This page intentionally left blank.

# 1. Introduction

## 1.1 Background Information

The following material, provided by the Project Authority (PA), provides some background information for the present project.

"On April 1, 2005, **D**efence **R**esearch and **D**evelopment **C**anada (DRDC) started a new applied research project in the **M**aritime **D**omain **A**wareness (MDA) Thrust: Information visualization and management for enhanced domain awareness in maritime security. This is a 4-year R&D project with the goal of enhancing the "maritime picture" through improved quality of information and novel, adaptive ways of visualizing that information. Part of this project is focused on visualization design and experimentation. The DRDC team wants to investigate the best way to visualize particular abstract concepts and then test to see if visualizing these abstract concepts can help improve understanding of the Recognized Maritime Picture (RMP), decision making based on the RMP and the efficiency of the RMP operators' duties. The work proposed under this call-up is exclusively related to testing the best way to visualize certain aspects of contact data and a surveyed environment."

The RMP is both a visual representation of surface contacts in the area of interest (AOI) on each coast extending out from Canadian coastal waters together with a database of metadata that is associated with each contact. The metadata comprises fields of information concerning attributes of the contact, such as location, speed, direction, name, flag, and a unique nine-digit identification number, known as a Maritime Mobile Service Identity (MMSI). Each contact may have various degrees of metadata associated with it depending upon the specific source or sources that feed the RMP.

As indicated by the PA, this incomplete information provides several challenges to the operators:

"For example….

- the metadata provided by a source being grossly out of date, and/or wrong, and/or missing;
- a sensor often providing false contact reports, with no indication of when it is false or not;
- the time-latency associated with each contact, which is different for every source;
- and contacts not being present, which can make it unclear whether there is an absence of a contact or the area hasn't been surveyed recently.

Similar challenges are faced by users attempting to understand the situation represented by the RMP. In addition, there is a further potential problem that the user and operator may not be cognizant of, namely, the sensor coverage or information coverage area varying due to weather, altitude (of a plane), ships not reporting as they should, or time of day."

It is possible that even experienced operators or users of the RMP may believe that contacts on a map are really in the positions they're displayed in. For example two contacts being shown next to each other may in fact be very far apart, as the positional information for the contacts are both time-late and not necessarily by the same amount of time. It is also possible that the belief exists that the contact report information for each contact is equally credible, and where there is an absence of a contact, no vessel exists.

Given this "uncertainty" in the metadata, the thrust of the present work is to identify ways in which such uncertainty could be readily brought to the operator's attention. To access such information in the current system would require significant overhead by the operator. For example, to determine what is known on a contact, the operator has to first select the contact and then scan a screen of data fields. Depending on what is shown, the operator may then have to review the track history of the contact or consult other data sources. Such an approach is time consuming and could not be realistically accomplished for the many thousands of contacts that may be present in the RMP at any time. Therefore, the provision of a readily comprehensible iconic representation of the uncertainty on a contact, would not only save countless hours of labor (should the task be undertaken for a subset of contacts), but would provide a new functionality that would greatly increase situation awareness concerning the precision of information provided by the RMP.

Thus the objective of the present work, as outlined in the Statement of Work (SOW) is:

*"to design and conduct an experiment to identify a pertinent set of visualization techniques that can potentially be used to enhance the current way of presenting the RMP. The specific objective is to design and conduct an experiment to evaluate the effectiveness with which an agreed upon set of visualization techniques represent the uncertainty inherent in the information presented in the RMP. In particular, the information we are interested in visualizing are:*

1. *Temporal uncertainty of the most current information on each track.*
2. *Position uncertainty of the most current information on each track.*
3. *Identity uncertainty of the most current information on each track.*
4. *Individual and cumulative coverage uncertainty of the sensors and information sources.[3]"*

## 1.2 Specific tasking

The following primary tasks have been summarized from the SOW

1. Familiarization with the problem by reviewing relevant literature
2. Selection of visualization techniques
3. Determine an appropriate environment to evaluate visualization candidates
4. Design, conduct and analyze an appropriate experiment
5. Provide a final report

[3] Subsequently, and as a result of continuing discussions with DRDC concerning project scope, it was agreed that the requirement to represent cumulative uncertainty would not be pursued.

# 2. Summary Literature Review

The first step in the project (as outlined in SOW Task 2.2a) was to conduct a literature review. The objective of the literature review was to familiarize the project team with either established or emerging concepts of information visualization and also to provide familiarization with the operational environment that the visualizations are relevant to. The PA provided the list of 13 documents to be reviewed (see Annex A for the full list). No database searches were conducted to find additional literature.

## 2.1 Defining and Categorizing Visualizations and Uncertainty

It became clear from the literature review that emerging concepts in the field of information visualization offer highly innovative ways of displaying complex data that are often characterized by a great deal of uncertainty (Celidnik & Rheingans, 2000). This is particularly true in the case of scientific data, where some degree of uncertainty is expected, and is part of the natural order of things. The efforts of developers and researchers within the area of data visualization are frequently predicated on the assumption that their world is essentially invisible or non physical in form and that their task is to provide a means of turning the ephemeral into something which is almost tangible, and which can be viewed, manipulated, or in which changes may be observed over time. Visualization is, in essence, the means by which phenomena previously unobservable become physically observable through the use of sophisticated graphic techniques.

The literature revealed a wide variety of types of uncertainty that had been subject to visualization techniques. Xie et al (2006) reflected upon different types of uncertainty and specifies three distinct granularities. The first granularity is at the individual data value level, where there is general uncertainty about each entry's quality. The second level of uncertainty is associated with a certain dimension (e.g. uncooperative participants in a survey may generate entire records of low confidence) and the final level of granularity of uncertainty is with the complete record (e.g. a defective sensor might make a specified attribute highly unreliable in the whole dataset). The focus of the present contract is on generalized data (and not real data) so the extent to which these granularities affect the data is not relevant. However, future projects may want to pay specific attention to these varying levels to ensure uncertainty models and visualizations are as representative and accurate as possible.

Different types of visualizations of uncertainty were outlined in Pang et al., 1997; Griethe & Schumann, 2006; Riveiro, 2007. The four main categories were:

1. Free graphical variables (e.g. color, size, position, focus, clarity, fuzziness, saturation, transparency and edge crispness) can be used to *alter* aspects of the visualizations to communicate uncertainty.
2. Additional static objects (e.g. labels, images or glyphs) can be *added* to the visualizations to communicate uncertainty.
3. Animation can be incorporated into the visualizations, where uncertainty is mapped to animation parameters (e.g. speed, duration, motion blur, range or extent of motion).
4. Uncertainty can be discovered by mouse interaction (e.g. mouse-over).

One of the first two methods (Altering or Adding) to express uncertainty was deemed by the project team and DRDC to be the desired approach for this project, as the goal was to have a static visual cue that requires no operator interaction.

## 2.2 Altering or Adding

A number of 'altering' type methods were encountered in the literature. Examples included the use of brightness, amplitude modulated distortion, etc (Celidnik & Rheingans, 2000). These techniques were reviewed as were the use of colour, grayscale, transparency, fuzziness and the use of rendering as a means to imply data quality and certainty (Stephenson et al, 2006). This information was of some value later in the present project when the team considered ways in which to convey uncertainty information about sensor coverage. At the forefront of this visualization study particularly with respect to ocean mapping is Colin Ware, who has produced numerous articles on the use of colour, texture, and motion in the design of 3D displays (e.g. Ware, 2004). As an alternative to changing the visualization specifically, another method involves 'adding' to the visualizations through new glyphs, grids, and other symbols that represent uncertainty information. Different applications require different types of additions, and a variety was explored in the literature. One example (Clausner and Fox, 2005) added symbology to timelines in order to clarify uncertainty in temporal information.

Ultimately, it was decided in conjunction with DRDC that *altering* the current RMP icons in order to include uncertainty information was not feasible. As such, the logical way forward for this project was to add a new 'symbol' to the icon that coded the desired uncertainty information.

## 2.3 Operational Environment

Two papers specifically addressed issues in the operational environment of the users and operators who work with the RMP. These papers provided insight into constraints of the RMP as well as the general tasks and goals associated with the RMP. The DRDC Atlantic Report CR 2006-038, (Davenport and Franklin, 2006) was especially useful in understanding the operational needs and future role and demands for the RMP. The Information Visualization contract report for DRDC Atlantic CR 2006-122 (Davenport and Risley, 2006) provided an invaluable insight into the scope and capability of emerging visualization techniques. While the latter provided some innovative approaches, it became quickly apparent that there would be difficulties in incorporating such techniques into the current system.

## 2.4 Final Remarks about the Literature

Attention was paid to the results of previous studies that had been conducted that outlined icon comparisons (Unger-Campbell and Baker, 2003) as well as which characteristics of colour seemed to work (e.g. saturation) and not work (e.g. hue) as an easily interpretable mapping of uncertainty (Riveiro, 2007; Cedilnik & Rheingans, 2000; Xie et al, 2006).

Finally, although the literature review helped to define and frame the possible types of visualizations of uncertainty, there was precious little discovered within the review that provided direction in terms of generating a solution. In particular, it was discouraging to find that no clear and comprehensive Human Factors guidelines had been proposed that would provide direction to finding design solutions for uncertainty visualization with respect to general principles.

# 3. Development Process for Visualization Concepts

In general, the process of developing the concepts involved frequent and detailed discussions between the HSI® team and DRDC. These discussions involved considerations of design constraints, criteria for evaluating proposed solutions and evaluation of proposed options.

## 3.1 Design Constraints

During the course of the literature review it became increasingly apparent that the requirements for development of visualization display concepts were constrained by the operational context and the need to maintain consistency with pre-existing RMP standards and symbology. In addition we were aware of the need to ensure that any proposed concepts must work within the task structure, tools and interfaces available to the operator.

Without doubt one of the major challenges in the search for appropriate visualization concepts was coming to terms with the constraints imposed by the context of use, and the need for such concepts to be accommodated within existing systems. Given that existing displays have a limited amount of 'real estate' in terms of screen size and the existence of potentially many hundreds of vessels of interest within a comparatively small area of ocean required that any proposed solutions needed to be compact, yet still capable of presenting the required levels of uncertainty information to the operator.

### 3.1.1 Role of the Indicator

There was general agreement reached that the role of the visualization icon would be to provide a cue or indicator concerning the quality of the RMP data on a contact, but the icon would not itself contain any data. This cue would allow operators to rapidly search and find contacts with data limitations (uncertainty), which in turn would then allow them to drill down into the associated metadata to determine the exact nature of the uncertainty, or lack of information.

The general concept of use for the icon was that it would be an "overlay" that could be brought up on the RMP as required by operators but would not be continuously displayed.

A series of discussions took place with DRDC that led to the decision that 3 levels of coding would be applied to each of the three dimensions of target uncertainty (IDENTITY, LOCATION, and TIME LATENESS). Subsequently, it was decided that just two levels of coding would be required to indicate sensor quality and sensor time lateness for area swaths depicting sensor coverage.

### 3.1.2 Physical Constraints

Without doubt, the most constraining aspect of the design solution was the need to ensure that any proposed design could be "attached" to the existing RMP contact icon in a manner that was consistent with the scale of the existing representation. This ruled out many of the visualization techniques proposed in the literature, where "design space" was frequently not a limiting factor.

Discussions suggested that actual pixel matrix of the design solution should be approximately either the same height or width of the contact icon, but should be smaller in total area than the icon.

This limited space meant that the use of detailed graphical representations (e.g. clocks and other meaningful graphics) could not be implemented within such a small area and that more abstract forms of representation (colour, shape, size) would be required.

### 3.1.3 Psychological Considerations

One of the major considerations was the need to ensure that there would be minimal ambiguity between the representations in terms of the uncertainty dimensions. This meant that representations should be visually and cognitively clearly separable. Therefore, for example, if colour were used, it should only be employed for a single dimension.

Further it was agreed that with respect to properties such as size, length, height or intensity that a *larger* symbol would always indicate *greater* uncertainty. It was agreed quickly that increasing the size of graphic representation would map readily to increased spatial uncertainty (analogous to the way "further on" circles are drawn around contacts in the current RMP to indicate approximate position). In contrast, there was more discussion concerning how to represent time uncertainty, with the literature not providing any clear guidance. Redefining time uncertainty as greater or lesser time lateness, helped to clarify the thinking about how this property could be appropriately represented.

### 3.1.4 Criteria for Evaluating Visualization Candidates for Coding Uncertainty

The following criteria were developed with DRDC for establishing some design boundaries for the proposed solutions.

1. There must be no "unintended interaction" in the representation when multiple types of uncertainty are displayed. That is, there must be a coherent and integrated solution so that the discriminability[4] and meaning of each coding level for each uncertainty source remains unaffected by the display of other representations of uncertainty. This will be evaluated by the team in developing prototype candidates.
2. Each uncertainty level must be uniquely understood by the operator and not rely on the presence of other uncertainty levels requiring a relative judgment. This means that there must be adequate "psychological distance" between the levels. This will be determined in pilot testing among the team.
3. The visual characteristics will maintain their discriminability under a range of ambient lighting conditions.
4. The use of colors and symbols will not conflict with any existing or potential systems for representing RMP symbology (e.g. NATO standard symbology).
5. The representation should be instantly perceivable and not rely on additional actions by the operator (e.g. hovering over a contact, right clicking etc)
6. The size of symbology should be appropriate for the zoom level.

---

[4] Discriminability: the ability of observer to distinguish symbols/text from the background and from each other.

7. The meaning of the coding should be directly perceivable (after suitable training) and not require the operator to rely on long term memory or consult supplementary information sources.
8. Wherever possible the coding should conform to existing psychological constructs and stereotypes (e.g. brighter=more, bigger=more, etc). However, it is not clear from the literature whether a higher stimulus value should be mapped positively on to the level of certainty or **un**certainty. To resolve this problem, we looked for clues among existing practices in the RMP and other tactical naval plots. In these cases, it was noted that a contact's positional uncertainty is generally indicated by the diameter of a circle centered on it. Where larger circles indicate a greater area in which the contact may be actually located. Therefore, not wishing to develop a design that would run counter to this existing population stereotype, it was decided that in order to represent spatial uncertainty and time lateness uncertainty, a larger stimulus value would correspond to greater **un**certainty.

   In the case of color application for coding, where strong populations stereotypes are in place for green, yellow and red, it was decided to adopt colors that would be consistent with this mapping. However, pure red was to be avoided because of the need to avoid its existing use as a warning indicator in the RMP (e.g. for hostile tracks). Therefore, colors such in the magenta-purple area of the spectrum were considered good alternatives as they maintained the notion of "redness" without the alerting properties of pure red.
9. Simpler symbol designs are preferred over more complex designs (where complexity can refer to aspects such as the number of external and internal contours, angular variability, matrix grain, number of turns/angles).
10. Use of alphanumeric characters is less preferred because of potential confusion with existing track data tags.
11. Blinking/flashing encoding on symbology should be not be used to avoid "Christmas lights" effect.
12. Use of blurring of symbols (e.g. Finger and Bisantz) should be avoided. Blurring can cause the eye focusing mechanism to hunt for a focusing solution. Over an extended viewing time there is a reasonable expectation that this could induce some unwanted visual symptoms.

## 3.2 Candidates Considered

### 3.2.1 Representation of Contact Uncertainty

Approximately 18 possible design options (and variants) for contact uncertainty were submitted to DRDC for review (see Annex B for examples). These were discussed extensively with DRDC and summarized in minutes of the February 27 and March 7, 2008 teleconferences. Ultimately, the most preferred options were circulated amongst the group and each individual provided preferences and comments on which candidates would be most suitable. In this way a consensus was reached on the final design options to evaluate in the experiment, and specific details on their design and implementation were then left to HSI® to develop the appropriate software specifications.

### 3.2.2 Representation of Sensor Uncertainty

The design alternatives for sensor uncertainty were considered to be somewhat more limited. It was agreed that the sensor quality would be represented by area fill which would be either different levels of a solid colour or different types of cross-hatching. Sensor time lateness would be represented by the border of the area fill, where line type (solid, broken, dotted), line width or line colour might be appropriate design options.

Samples of possible design options were shared with the team and relative merits and limitations were discussed. With respect to area fill it was agreed that the design chosen should (i) provide adequate conspicuity from the RMP background, (ii) allow any displayed symbols to be easily discriminated when located on the swath (iii) allow two levels to be readily comprehended and (iv) not be so conspicuous as to unbalance the overall appearance of the RMP. In considering the area boundary some constraints considered were potential representational problems with broken lines or dots, the need for simplicity, the need to keep borders narrow. Further the use of colour in the border should intuitively map onto time lateness.

## 3.3 Final Candidates

### 3.3.1 Contact Uncertainty

Two sets of symbols were designed for use in the first phase experiment. The first of these was the 'Rectangle and Circle' design. This design is outlined first. The second was the 'Lego'[5] design which will be outlined subsequently. The choice of colors and terminology for each level of uncertainty used in these designs was reached through group consensus with the goal of meeting the needs of the planned experimental study. It was noted that if such designs were actually adopted for integration into the RMP there would need to be further exploration on what would be the most appropriate colour palette and phrasing for the operational environment. In addition, it would be necessary to explore the extent of uncertainty each level represents.

Identity uncertainty is represented by the background colour using common population stereotypes based upon traffic signals, with the exception that red was replaced by "plum" to avoid any unwanted alerting associations and to avoid conflict with any existing red coding in the RMP. Green indicates 'Good' identity information, 'Amber' = poor identity information, and 'plum' indicates that 'no identity information' is available.

Spatial uncertainty was represented by a black filled circle where the size of the circle indicates the precision of the information available. The smallest circle indicates "precise" location, the intermediate size "approximate position" and the largest indicating an 'unknown' spatial position.

Time information is represented by the position of the circle in the background rectangle. Current time is indicated by a leftmost position, "old" time by a central position and "out of date" by the rightmost position. Essentially, the position indicates the "time lag" between now and when the information was received, so moving to the right means moving down the timeline.

Similar principles were used in the 'Lego' 'L' shaped design. Identity was again represented by colour as with the 'Rectangle and Circle' design, although some amendments were made to the colour scheme following feedback from DRDC. The original proposal was to use the same colors

[5] This was a label of convenience based upon the similarity of the design to LEGO building blocks and is not associated with the registered trademark of The LEGO Group.

as the Rectangle design, but when the original icons were shrunk down to the appropriate scale several potential problems were observed necessitating some ”tweaking” of the original colors. The final decision was to use the same ‘Green’ for ‘Good’ Identity information, but to use ‘Yellow’ and ‘Orange’ to illustrate ‘Poor’ and ‘No’ Identity information respectively.

Spatial uncertainty was represented by the number of squares in the horizontal bar. Since more squares meant “greater” uncertainty, three squares were used to indicate unknown position, two squares to indicate approximate position and a single square to indicate precise position.

Time uncertainty was indicated by the number of vertical squares, where again more squares meant a longer time gap, with a single square at the bottom indicating the time information is current, two indicating the time information is old and three that it is out of date.

Overall, this design maps intuitively in the following way. When the Lego design is compact (two single vertical and horizontal blocks) time and position uncertainty is low, when it more extensive, (three blocks in each dimension) time and position uncertainty are high.

Both the horizontal and vertical bars are composed of three segments. Highest levels of ‘certainty’ are represented by a single colored segment (lowest of the three in the vertical bar) representing ‘Time’ information, and the leftmost segment of the horizontal bar representing ‘Current’ time information.

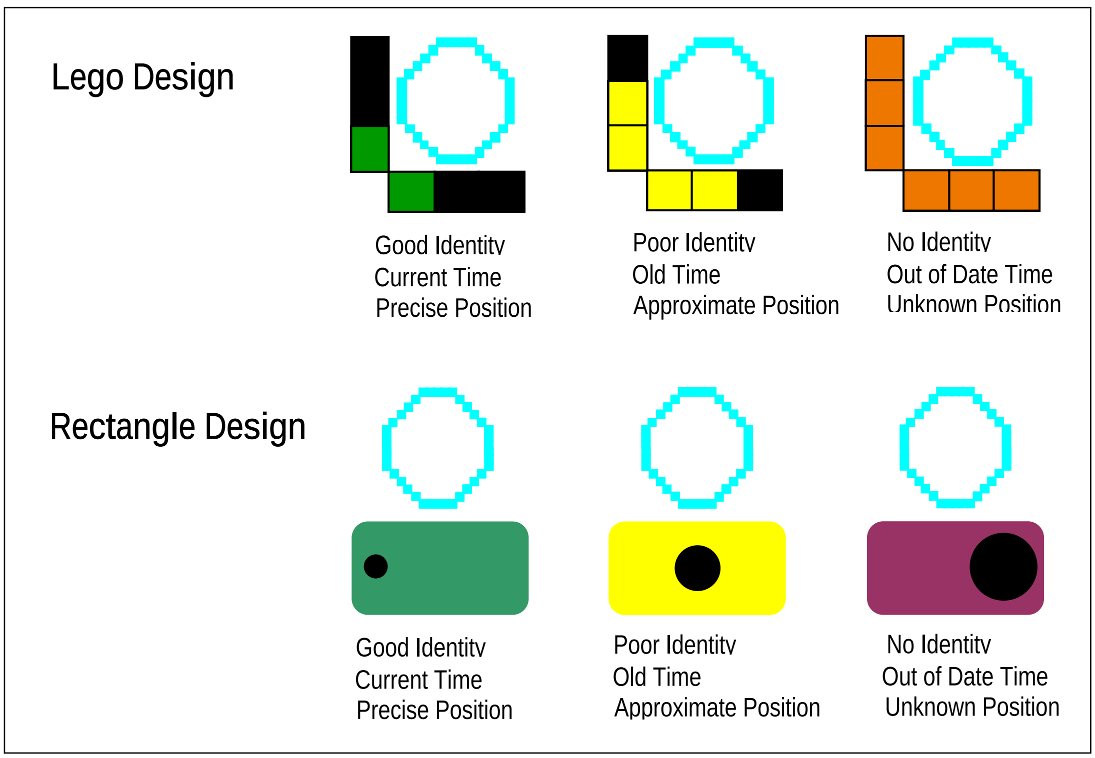



**Figure 1: Samples of the two design concepts**

### 3.3.2 Sensor Quality and Time Lateness

The two design concepts to represent sensor quality were a grey solid fill or cross hatching. Examples of the two levels of these to indicate good or poor quality sensor coverage are shown in the next figure.  In addition, sensor time lateness was indicated by either the type of the swath border (solid/broken) or the border colour (green/gold).  Note that the blue background was adopted for use in the experiment since this corresponded with the way in which the RMP is currently presented.  It was noted that based upon the existing Human Factors literature such a choice of blue for the background is not considered desirable or optimal for visually intensive search tasks

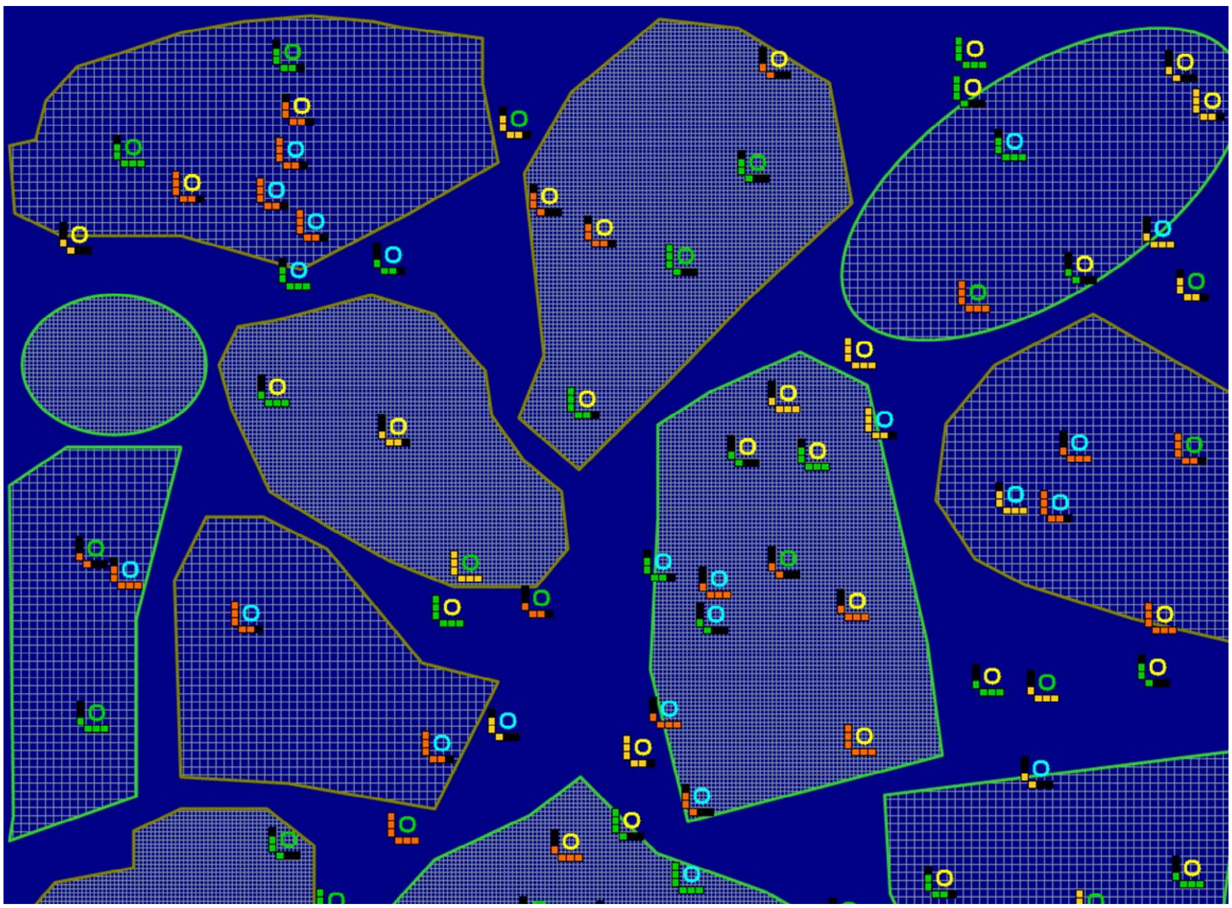


**Figure 2: Example screen display with grid and colored border swaths (not to scale) and Lego targets.**

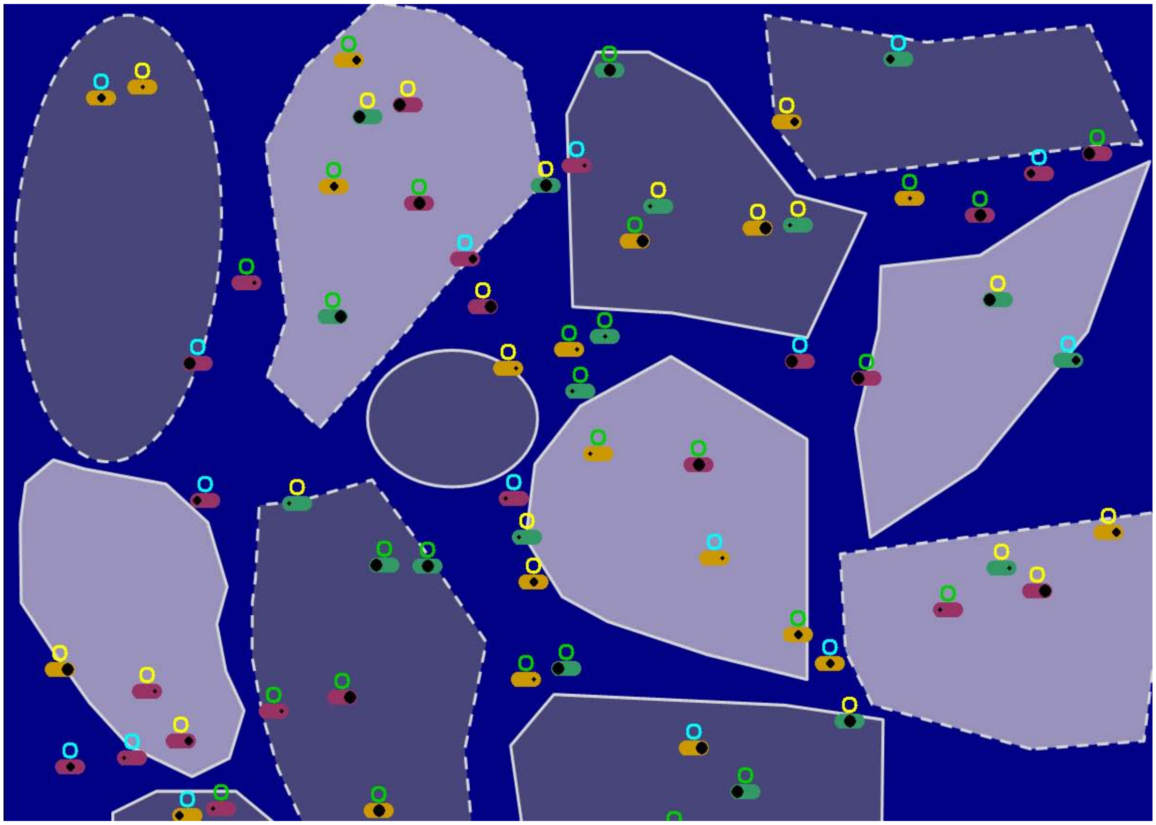


**Figure 3: Example screen display with shading and solid/dashed border swaths (not to scale) and Rectangle targets.**

# 4. Experiment Development

## 4.1 Selection of a Test Environment

In developing the concept of evaluating visualization design options, the PA had initially considered the following possibilities for collecting quantitative evaluation data:

- Story boards
- A simulation environment available at DRDC Atlantic
- A stand alone experimental environment (e.g. conduct testing at Humansystems Guelph facility)

A brief discussion was conducted with the PA concerning the advantages and disadvantages of each environment, which are summarized in the table below. The conclusion was quickly reached that the best option would be to proceed with testing at the Humansystems office facility.

**Table 1: Evaluation of potential test environments**

| Option | Advantages | Disadvantages |
|---|---|---|
| Storyboards | • Low time and cost overhead to produce<br>• Portable | • Limited functionality<br>• Low face validity of representations<br>• Only subjective evaluations possible |
| DRDC Simulation Lab | • Hardware available<br>• Ability to represent high face validity displays<br>• Potential ability to collect human in the loop, quantitative performance data | • Questionable availability of test subjects<br>• High logistical overhead for contractor to implement experiment |
| Humansystems Test Lab | • Hardware available<br>• Ability to represent high face validity displays<br>• Potential ability to collect human in the loop performance data<br>• Local software development resources available<br>• Local availability of assistants to administer experiment<br>• Available subject pool | |

## 4.2 Development of the Experimental Paradigm

The development of the paradigm was guided and constrained by a number of factors that are outlined below:

1. Face validity: the experimental display and the subject's task should be representative of actual tasks that might be performed by actual operators on the RMP graphic representation
2. The display will not be dynamic as this would impose significant demands on the software development
3. The task should allow the collection of real time performance data such as response time and accuracy
4. Since access to actual RMP operators would not be feasible, the task should allow untrained subjects to reach proficiency with minimal training
5. Symbols used to represent uncertainty should be compatible with the existing RMP symbology, at approximately the same scale and resolution
6. Three levels of uncertainty for each uncertainty dimension would need to be uniquely represented in the icon.
7. "Sensor coverage" would be represented by an area fill on which target icons were superimposed. Two levels of sensor quality and two levels of sensor time lateness would need to be represented.
8. The task should allow the collection of a statistically reliable data set in a subject test session lasting between 45-75 minutes

Resolution of options for the test paradigm was achieved through a number of discussions with the PA and evaluation of the merits of various approaches. Subsequently, agreement was reached on a paradigm, which is outlined in the next section.

### 4.2.1 Overview of the Paradigm

Subjects were presented with a display which contained a number of tracks each represented by a facsimile of the existing RMP symbology for surface contacts. Attached to each piece of symbology was an icon that represented a level of uncertainty for the track identity, time lateness and spatial position. Each display typically contained approximately 60+/-5 symbols (phase I) or 80+/-5 symbols (phase II). In the second phase of the study, 12 areas representing sensor coverage were added to the display; for each area the quality (area fill) and recency (area border) of the sensor coverage were depicted.

Subjects were trained using a prepared PowerPoint presentation to become familiar with how each type of uncertainty and each level was coded in the iconic representation, as well as the coding of quality and recency for the sensor coverage (when applicable; phase II). Following training, subjects participated in test sessions containing a number of trials. Each trial comprised a written description of the search criterion prior to presentation of the test screen. For example, "find all targets with no identity information, out-of date time information that fall within good quality, current sensor coverage". Subjects then searched through the test screen, clicked on targets that met the criteria and the software recorded the elapsed time to select each target. To ensure some reasonable level of task difficulty and to allow the potential for selection errors, the display would comprise a mix of targets that met the criteria among a background of other target symbols that did not meet the criteria. The software also recorded the time required by subjects to read and comprehend the instructions prior to starting the trial, as well as the time between the last target selected and the subject clicking the “done” button to indicate that there were no more targets to be found. To avoid potential confusion of subjects not remembering which targets they had selected, once a target had been clicked it would be dimmed on the display.

Complete specification of the final experiment design can be found in sections 4.4-4.7.

## 4.3 Test Participants

### 4.3.1 Ethics Approval

As required for all DRDC studies involving human participants, the proposed testing protocol required approval by the Human Research Ethics Committee (HREC), which is based at DRDC Toronto. Accordingly, following the HREC standard forms and guidelines, a protocol was prepared by HSI® and reviewed by DRDC prior to submission to the HREC. This protocol was subsequently reviewed by HREC on February 13, 2008, with Michael Matthews from HSI® and Sharon McFadden, representing the PA, in attendance. Following questioning from the committee and subsequent written direction from the committee for certain changes, a revised protocol was approved. This protocol is provided in Annex C.

### 4.3.2 Recruitment of Test Participants

Test participants, or subjects, were recruited from a pool of individuals in the Guelph area who had previously participated in studies at HSI® and had indicated a willingness to be involved in future studies. Participants were contacted by telephone and explained the nature of the study as prescribed in the protocol. In particular, the need for them to be available for several sessions of testing was stressed.

As a result, twenty five subjects were identified for potential participation. Subsequently, thirteen participated in phase I of the study and eleven were available to continue in phase II. A preliminary analysis of the data indicated that the sample size was sufficiently large to ensure statistical reliability and sensitivity to experimental manipulations. An additional six subjects participated in a pilot study during the development of phase I. One of the 13 subjects (female) participated in phase I, but dropped out part way through phase II (therefore her data was excluded from the analysis). The final sample comprised 6 females and 5 males who were required to self report whether they had any problems with their vision and whether they had normal colour vision.

Subjects were paid $22 for each test session in both phases I and II. The initial session in phase II often required more than 2 hours of participation due to expanded training and therefore when necessary, subjects were paid $27.

### 4.3.3 Equipment

Two test stations were set up to allow for parallel test sessions where feasible; each set-up comprised the same hardware. The test software was run on a portable DELL notebook and the display was presented on an Acer 19” LCD monitor. This was calibrated using a Minolta CS100 color photometer to ensure that the two systems used for parallel testing of subjects provided equivalent images. The two test rooms were set up to provide approximately equivalent ambient illumination, which was 150Lux measured at the keyboard. This corresponds to moderately low level office lighting. Each notebook had an optical mouse attached which enabled subject responses to be collected.

## 4.4 Stimuli

### 4.4.1 Contact (track) Icons

Three levels of uncertainty (low, medium, high) were coded for each of three dimensions of contact identity, time lateness and spatial position.

The verbal labels for the uncertainty levels were:

- **Identity:** *good, poor, none*
- **Time:** *current, old, out of date*
- **Spatial location**: *precise, approximate, unknown*

Samples of both designs and uncertainty level descriptors are shown in figure 1 in the previous section.

The photometric properties[6] of the stimuli are shown in Table 2.

**Table 2: Target Stimulus properties**

| Rectangle | Luminance (cd/m$^2$) | x | y |
|---|---|---|---|
| Blue background | 12 | .257 | .224 |
| Identity - green | 42 | .29 | .44 |
| Identity - gold | 57 | .45 | .437 |
| Identity - mauve | 23 | .427 | .309 |
| Lego | | | |
| Blue background | 12 | .257 | .224 |
| Identity - green | 53 | .266 | .466 |
| Identity - yellow/sand | 79 | .453 | .408 |
| Identity - rose/orange | 53 | .507 | .395 |

The “circular” icon representing the target was randomly assigned the colour cyan, yellow or green.

[6] Measured with a Minolta CS100A, using the Yxy co-ordinate system from a distance of approximately 20cm using a close up lens. The test stimuli were measured in the centre of the screen.

### 4.4.2 Swath Areas

Two types and two levels of uncertainty were represented in the swath areas: quality of the coverage (strong, weak) and time lateness (current, out of date). Two design options for representing these aspects of uncertainty were presented as shown in the following figures. Quality coverage was either represented by the coarseness of a grid fill (strong= fine grid, weak=coarse grid) or by grey area fill (strong=light grey, weak=dark grey) at two levels. Time lateness coverage was either represented by the colour of the border (current=green, out-of-date=gold) or by the continuity of the border (current=solid, out-of-date=broken).

**Table 3: Swath properties**

| | Quality- Strong | Quality- Weak | Time- Current | Time- Out of Date |
|---|---|---|---|---|
| Design 1 | Fine grid<br>(8 lines/cm) | Coarse Grid<br>(4 lines /cm) | Green Border<br>L=58 cd/m$^2$,<br>x=.272, y=.579 | Gold Border<br>L=99 cd/m$^2$,<br>x=.443, y=.477 |
| Design 2 | Light gray<br>L=56 cd/m$^2$,<br>x=.272, y=.274 | Dark Gray<br>L=12 cd/m$^2$,<br>x=.232, y=.205 | Solid White Border | Broken White Border |

Examples of each of the designs with target symbols superimposed are shown in figures 2 and 3 in the previous section.

## 4.5 Software

The software to present the stimuli and collect response data was coded in PHP with some embedded Javascript and designed to run as a web page. To run the software, an Apache webserver that supports PHP version 5 or higher is required. The source code for the experiment has been provided to the PA separate to this document.

The delivered software was tested by having analysts run through each condition and note their selections. Some deliberate mistakes were also made. Then, the data record was checked against the notes for consistency. If errors were found, they were reported back to the software developer for remediation. This process iterated until the team was satisfied that the software was working correctly.

In the case of the software required to support the swath trials some additional steps were taken, as our routine testing indicated that the criterion contacts were not being distributed appropriately across all of the criterion swath areas. Theses steps involved creating functionality that would allow the analyst to preview a trial screen and then “regenerate” it until a desired target distribution was obtained. This screen was then saved to be run in the experiment. This procedure was conducted for all 40 trials of each test session for Phase II, and took considerable time to complete.

## 4.6 Design

The experiment comprised of two phases. In phase I targets were presented without any swaths present, hence subjects were only required to search for targets based upon target uncertainty criteria. In phase II of the experiment swaths were introduced, which required subjects to search for criterion contacts in criterion swath areas.

There were 63 combinations of target stimuli used in the experiment as shown in Table 4. Twenty seven of these represented all combinations of three contact uncertainty search criteria; twenty seven represented two criteria and nine represented a single criterion.

**Table 4: List of target combinations of search criteria**

| Number | Time | Spatial | Identity | Number | Time | Spatial | Identity | Number | Time | Spatial | Identity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Three criteria combinations* | | | | *Two criteria combinations* | | | | *Single criterion* | | |
| 1 | L | L | L | 28 | L | L | | 55 | L | | |
| 2 | L | L | M | 29 | L | M | | 56 | M | | |
| 3 | L | L | H | 30 | L | H | | 57 | H | | |
| 4 | L | M | L | 31 | M | L | | 58 | | L | |
| 5 | L | M | M | 32 | M | M | | 59 | | M | |
| 6 | L | M | H | 33 | M | H | | 60 | | H | |
| 7 | L | H | L | 34 | H | L | | 61 | | | L |
| 8 | L | H | M | 35 | H | M | | 62 | | | M |
| 9 | L | H | H | 36 | H | H | | 63 | | | H |
| 10 | M | L | L | 37 | L | | L | | | | |
| 11 | M | L | M | 38 | L | | M | | | | |
| 12 | M | L | H | 39 | L | | H | | | | |
| 13 | M | M | L | 40 | M | | L | | | | |
| 14 | M | M | M | 41 | M | | M | | | | |
| 15 | M | M | H | 42 | M | | H | | | | |
| 16 | M | H | L | 43 | H | | L | | | | |
| 17 | M | H | M | 44 | H | | M | | | | |
| 18 | M | H | H | 45 | H | | H | | | | |
| 19 | H | L | L | 46 | | L | L | | | | |
| 20 | H | L | M | 47 | | L | M | | | | |
| 21 | H | L | H | 48 | | L | H | | | | |
| 22 | H | M | L | 49 | | M | L | | | | |
| 23 | H | M | M | 50 | | M | M | | | | |
| 24 | H | M | H | 51 | | M | H | | | | |
| 25 | H | H | L | 52 | | H | L | | | | |
| 26 | H | H | M | 53 | | H | M | | | | |
| 27 | H | H | H | 54 | | H | H | | | | |

In phase I, for each combination, between 8-12 targets were presented on each trial in a background of 50 non-targets. All 63 combinations were presented in a random order to subjects. The two design concepts were run in different sessions with the order counterbalanced across subjects.

This phase, therefore, comprised a 2x3 completely within subject design, with two levels of variable (i) –design concept and three levels of variable (ii) – number of search criteria.

In Phase II of the experiment, a representative subset of 23 of the above target combinations was used (yellow cells), because it would have made the test sessions too lengthy, and the time required of subjects too great, if all 63 combinations were to be represented. On each trial, between 1 and 12 targets were present (average=6-7) and there were 40 trials within a trial block.

A summary of the design for phase II of the experiment is shown in the next table.

**Table 5: Phase II Summary of Test Sessions**

| Session | Target | Swath border | Swath Fill |
|---|---|---|---|
| 3A | LEGO | Green/Gold | Grid |
| 3B | LEGO | Green/Gold | Shading |
| 4A | LEGO | Solid/Dashed | Grid |
| 4B | LEGO | Solid/Dashed | Shading |
| 5A | RECTANGLE | Green/Gold | Grid |
| 5B | RECTANGLE | Green/Gold | Shading |
| 6A | RECTANGLE | Solid/Dashed | Grid |
| 6B | RECTANGLE | Solid/Dashed | Shading |

Each session was preceded by subject training using PowerPoint slides, followed by 3 practice trials. Within each session there were 10 trials for each of the four conditions of swath uncertainty (quality =strong or weak; time=current or out of date). These conditions were randomly presented.

In phase II the design comprised a 2x2x3x2 design with two border design concepts, two border fill concepts, three levels of target uncertainty, and two levels of target icon design. Subjects were counterbalanced across swath fill for each session (i.e., some started with 3A then 3B while others started with 3B and then 3A and so on).

## 4.7 Procedure

### 4.7.1 Phase I

Subjects completed a PowerPoint training session supervised by the experiment administrator (EA); this took approximately 20-30 minutes to complete. A sample training package is provided in Annex E. They then proceeded to the first trial of the experiment.

Each trial was preceded by written instructions, which indicated the search criteria for that trial. The wording of the criteria was as follows.

*Three criterion example*:

For this task, please select all targets that have:

- Out-of-date time information
- Approximate spatial position information
- No identity information

*Two criterion example:*

For this task, please select all targets that have:

- Current time information
- Approximate spatial position information

*Single criterion example:*

For this task, please select all targets that have:

- No identity information

Subjects were generally instructed to search carefully and that accuracy was more important than speed. Subjects were also informed that it was not possible to un-click a target once it had been selected. Subjects had access to printouts of the relevant icon designs and their associated criteria throughout the trials.

When the subjects were ready to proceed with the trial they pressed the "Start Trial" button. Subjects then searched through the screen and clicked on each target that they found. When the target was clicked, the intensity was reduced. Once the subjects were satisfied that all targets that met the criteria were found, they pressed the "Done" button. At any time during the trial, subjects could press the "Instructions" button to review the criteria for the trial. After the completion of each trial, the instructions were immediately presented for the following trial.

During the trial subjects had access to printouts of the relevant concept design showing the icon that matched each reference criterion.

A software timer recorded the following:

- Time from Instruction onset to pressing "Start Trial" button
- Time from above to each subsequent target selection
- Time from trial onset to "Done" button

In addition, the software recorded the accuracy of each target selection.

Following each session, subjects filled out a short questionnaire to evaluate their subjective impressions of various aspects of the design concept.

Each session took approximately 90-110 minutes to complete. At the end of the session subjects completed a short questionnaire that attempted to evaluate factors such as how easy it was to comprehend the coding and to discriminate the different levels, workload and visual discomfort. (see Annex D).

### 4.7.2 Phase II

Subjects completed 2 blocks of experimental trials in each of the four sessions in Phase II. Each session began with a PowerPoint training session supervised by the EA. For the first session (i.e., Session 3), training required approximately 30 minutes to complete. There were test questions throughout the training to ensure subjects had an understanding of the meaning of each swath display and the target icons, as well as the required search task. Subsequent sessions (i.e., 4 through 6) required between 5 and 10 minutes of training to remind subjects of the meaning of the swaths and targets. Subjects then proceeded to the practice session, which consisted of three trials. The EA remained with the subjects for these trials to ensure accuracy and to answer questions when necessary. Once complete, subjects began the first block of 40 experimental trials.

Each trial was preceded by written instructions, which indicated the search criteria for that trial.

The wording of the criteria was as follows.

*Three criterion example*:

For this task, please look in **strong** coverage that is **out-of-date** and select all targets that have:

- Current time information
- Precise spatial position information
- Good identity information

*Two criterion example:*

For this task, please look in **weak** coverage that is **current** and select all targets that have:

- Out-of-date time information
- Approximate spatial position information

*Single criterion example:*

For this task, please look in **weak** coverage that is **out-of-date** and select all targets that have:

- Old time information

Subjects were generally instructed to search carefully and that accuracy was more important than speed. Subjects were also informed that it was not possible to un-click a target once it had been selected. Similar to Phase I, subjects had access to printouts of the relevant concept design showing the icon that matched each reference criterion, and the meaning of the swath fill and border.

When the subjects were ready to proceed with a trial they pressed the "Start Trial" button. Subjects then searched through the screen and clicked on each target that they found fitting the specified criteria within the appropriate swath. When the target was clicked, its intensity was reduced. Once the subjects were satisfied that all targets that met the criteria were found, they pressed the "Done" button. At any time during the trial, subjects could press the "Instructions" button to review the criteria for the trial. After the completion of each trial, the instructions were immediately presented for the following trial.

A software timer recorded the following:

- Time from Instruction onset to pressing "Start Trial" button
- Time from above to each subsequent target selection
- Time from trial onset to "Done" button

In addition, the software recorded the accuracy of each target selection.

Following the 40 trials, subjects filled out a short questionnaire regarding the perceived workload of the task; this was followed by a short break. Subjects then proceeded to have a 2 to 5 minute training session where the EA presented several printouts showing the changes in the task for the next block of trials (whereby the swath fill was changed, but all other variables remained constant). After training, subjects completed three practice trials in the presence of the EA, followed by 40 experimental trials. At the completion of the second block of trials, subjects filled out another short workload questionnaire, as well as questionnaires regarding their subjective impressions of the swath fill and border (the border questions were only relevant at the end of sessions 4, 5 and 6 because subjects were only familiar with one type of border at the end of session 3).

Each session (i.e., 2 blocks) took approximately 115-135 minutes to complete.

## 4.8 Planned Statistical Analysis

To compare means among the test conditions Analysis of Variance (ANOVA) was used to assess main effects and any interactions. The criterion significance level was set at $p<.05$, which means any obtained differences could only have occurred by chance five times in one hundred.

The standard method for reporting the significance for the comparison in question is $F(a,b)=x$, $p<=y$, where a is the number of degrees of freedom for the means concerned, b represents the degrees of freedom for the ANOVA error term, x is the computed F value, p represents probability and y represents the actual probability value.

Where ANOVA resulted in significant main effects involving three or more means, or significant interactions, then individual means were compared using the Critical Difference (CD) test. This is computed as follows:

$$\mathbf{CD}=t*2*MSerror/n,$$

Where t=standardized t value for appropriate degrees of freedom (obtained from tables), *MSerror* is the appropriate mean square error term resulting from the ANOVA, and *n* is the number of subjects that each mean is based on. While multiple comparisons using this metric provide no protection against increasing the experiment rate for making a Type I error[7], they do ensure that small differences between means are identified, which is important in exploratory research of the present type.

[7] Incorrectly rejecting the null hypothesis when in fact there is no difference among means

# 5. Phase I Results: Target Search with No Sensor Coverage Overlay

## 5.1 Search Time

This is the mean time required to locate a single target in each individual test screen. For simplicity of presentation, the mean time across all trials for each level of the three combinations of target criteria (one, two or three) have been calculated and are presented in the following figure for the two design concepts. Error bars equivalent to one standard error (standard deviation/$n^{0.5}$) are indicated.

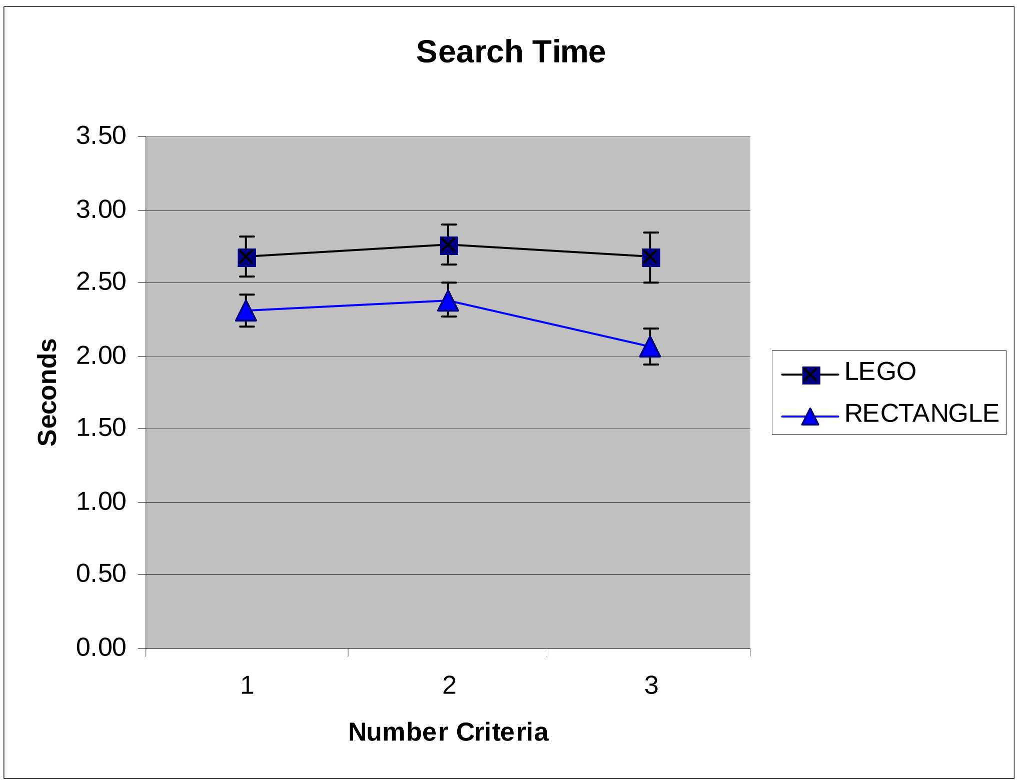



**Figure 4: Mean search time as a function of the number of search criteria**

An initial analysis was done of whether there was any effect of the test order of the design concept (i.e. whichever was presented first). This analysis showed no significant difference between the testing orders, therefore this variable was not included in subsequent analyses.

As can be seen from figure 4, the Rectangle design showed faster search times for all criterion levels. ANOVA showed significant main effects for type of design ($F(1,12)=553.3$, $p<=.001$) and number of search criteria ($F(2,24)=7.68$, $p<.001$), and an interaction between these variables ($F(2,24)=4.44$, $p=.02$).

Further exploration using t-tests (critical difference between means, $t(12)=.18$, $p<.01$) showed that the Lego condition was slower for all three criterion conditions. Within the Lego condition there was no difference among the three criterion levels. For the Rectangle condition, the three criterion search was significantly faster than the one or two criteria search, which did not differ from each other.

### 5.1.1 Individual Differences

The following figure shows mean search time for the two conditions for all subjects collapsed over all target combinations. As can be seen, search performance was faster in the Rectangle condition for all subjects, although the magnitude of the difference varied to some degree.

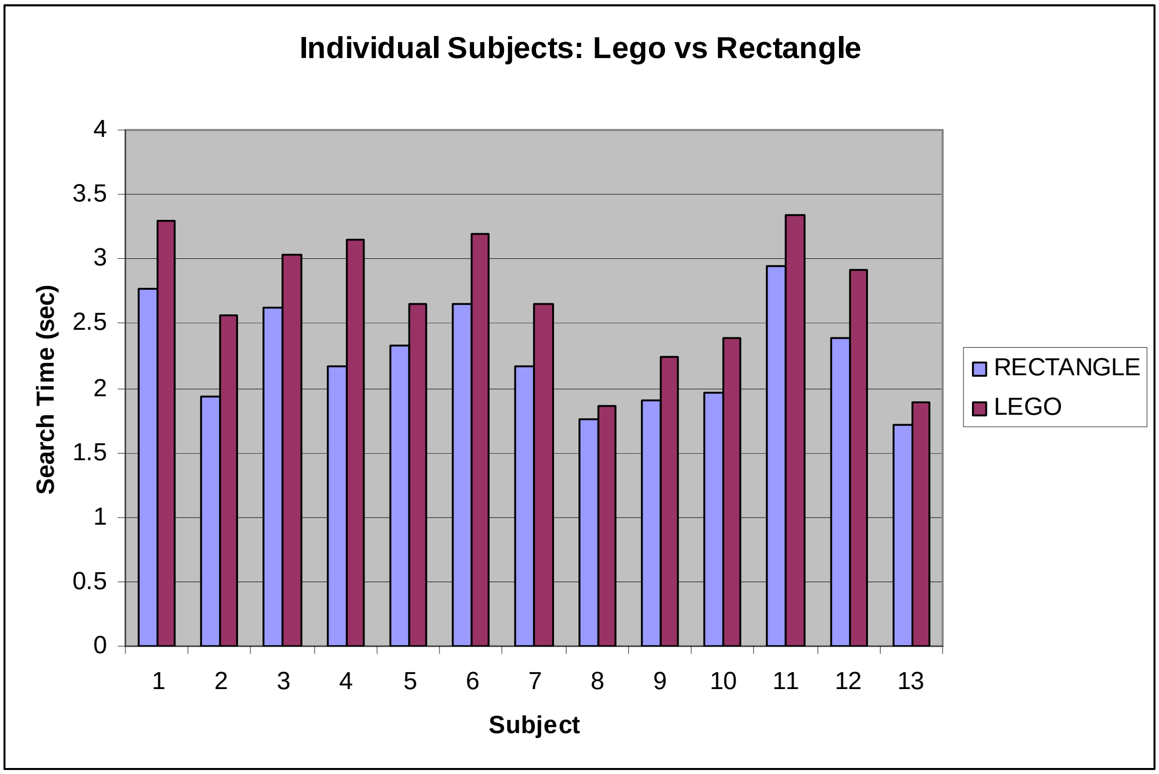



**Figure 5: Search time for individual subjects**

These data are broken down further in the next two figures which show individual results for the three levels of criteria.

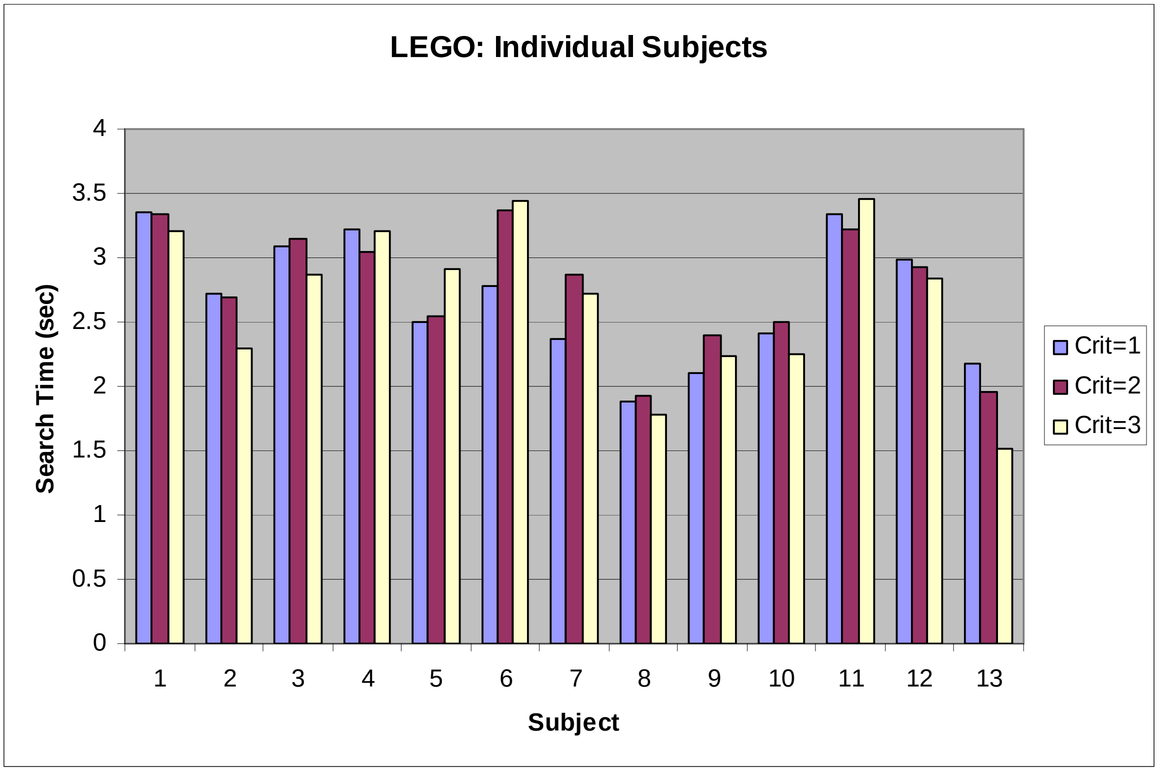



**Figure 6: Individual subjects: LEGO search time**

Comparing just search with one criterion versus three criteria for the Lego design, seven subjects showed faster performance with three criteria, but four showed a trend in the other direction and the remainder showed an inconclusive trend.

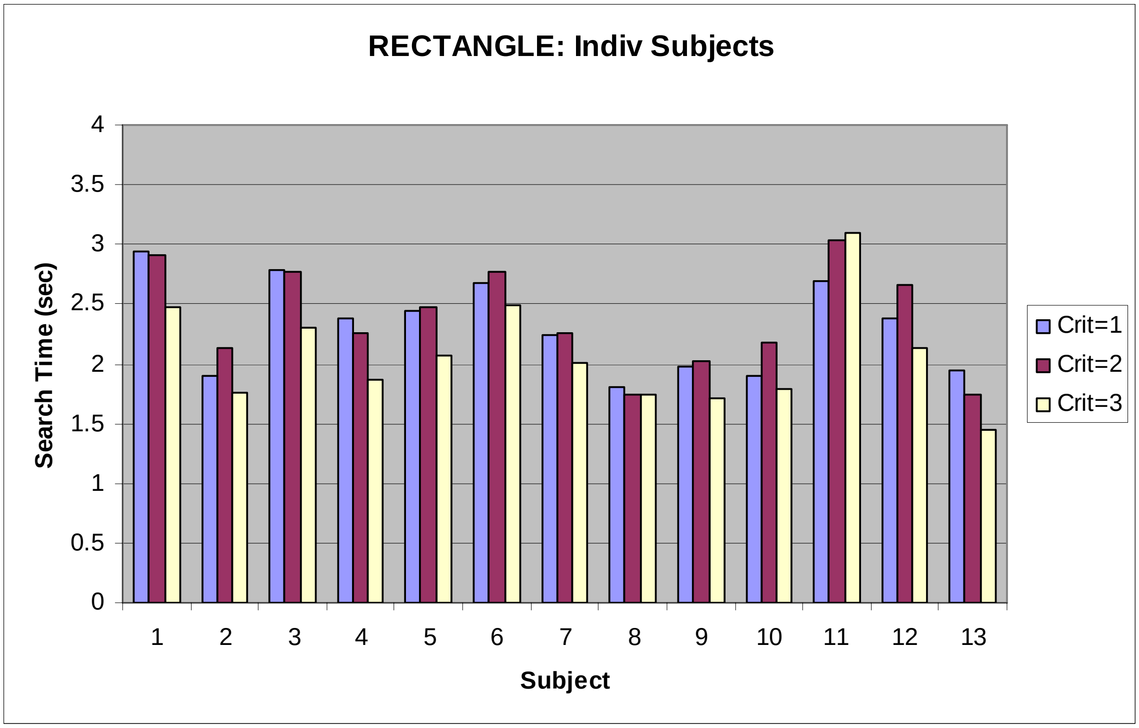



**Figure 7: Individual subjects: RECTANGLE search time**

Comparing just search with one criterion versus three criteria for the Rectangle design, eleven subjects showed faster performance with three criteria.

The following table shows the individual mean search times for each subject for each condition. Most noticeable among the individual differences were three subjects whose overall mean performance was approximately 20% slower than the group mean. These are highlighted in yellow in the table.

**Table 6: Individual subject mean search time**

| | LEGO: OVERALL MEAN=2.71 sec | | | | | RECTANGLE: OVERALL MEAN=2.22 sec | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Subject | 1 | 2 | 3 | MEAN | STDEV | 1 | 2 | 3 | MEAN | STDEV |
| 1 | 3.3 | 3.3 | 3.2 | 3.3 | 0.08 | 2.94 | 2.91 | 2.47 | 2.70 | 0.27 |
| 2 | 2.7 | 2.7 | 2.3 | 2.6 | 0.24 | 1.90 | 2.13 | 1.75 | 2.03 | 0.19 |
| 3 | 3.1 | 3.1 | 2.9 | 3.0 | 0.14 | 2.79 | 2.78 | 2.30 | 2.52 | 0.28 |
| 4 | 3.2 | 3.0 | 3.2 | 3.2 | 0.10 | 2.38 | 2.25 | 1.87 | 2.40 | 0.26 |
| 5 | 2.5 | 2.5 | 2.9 | 2.7 | 0.23 | 2.44 | 2.47 | 2.06 | 2.23 | 0.23 |
| 6 | 2.8 | 3.4 | 3.4 | 3.2 | 0.37 | 2.67 | 2.78 | 2.49 | 2.64 | 0.14 |
| 7 | 2.4 | 2.9 | 2.7 | 2.7 | 0.26 | 2.24 | 2.25 | 2.00 | 2.17 | 0.14 |
| 8 | 1.9 | 1.9 | 1.8 | 1.9 | 0.07 | 1.81 | 1.74 | 1.74 | 1.60 | 0.04 |
| 9 | 2.1 | 2.4 | 2.2 | 2.2 | 0.15 | 1.98 | 2.03 | 1.71 | 1.86 | 0.17 |
| 10 | 2.4 | 2.5 | 2.2 | 2.4 | 0.13 | 1.90 | 2.18 | 1.79 | 1.94 | 0.20 |
| 11 | 3.3 | 3.2 | 3.4 | 3.3 | 0.11 | 2.69 | 3.04 | 3.10 | 2.79 | 0.22 |
| 12 | 3.0 | 2.9 | 2.8 | 2.9 | 0.07 | 2.38 | 2.67 | 2.13 | 2.36 | 0.27 |
| 13 | 2.2 | 2.0 | 1.5 | 1.9 | 0.34 | 1.94 | 1.74 | 1.45 | 1.63 | 0.24 |
| STDEV | 0.49 | 0.48 | 0.62 | | 0.10 | 0.38 | 0.43 | 0.44 | | 0.07 |

In general, it appears that the Lego design produces more inter-subject variability than the Rectangle design.

### 5.1.2 Effect of Time on Task

The effect of practice on the task is shown in the next figure. The data represent the mean search time for successive blocks of trials (with each trial comprising between 8-12 targets). Of course, the specific target combination varies randomly across the trial sequence, therefore the data below are collapsed over these target combinations.

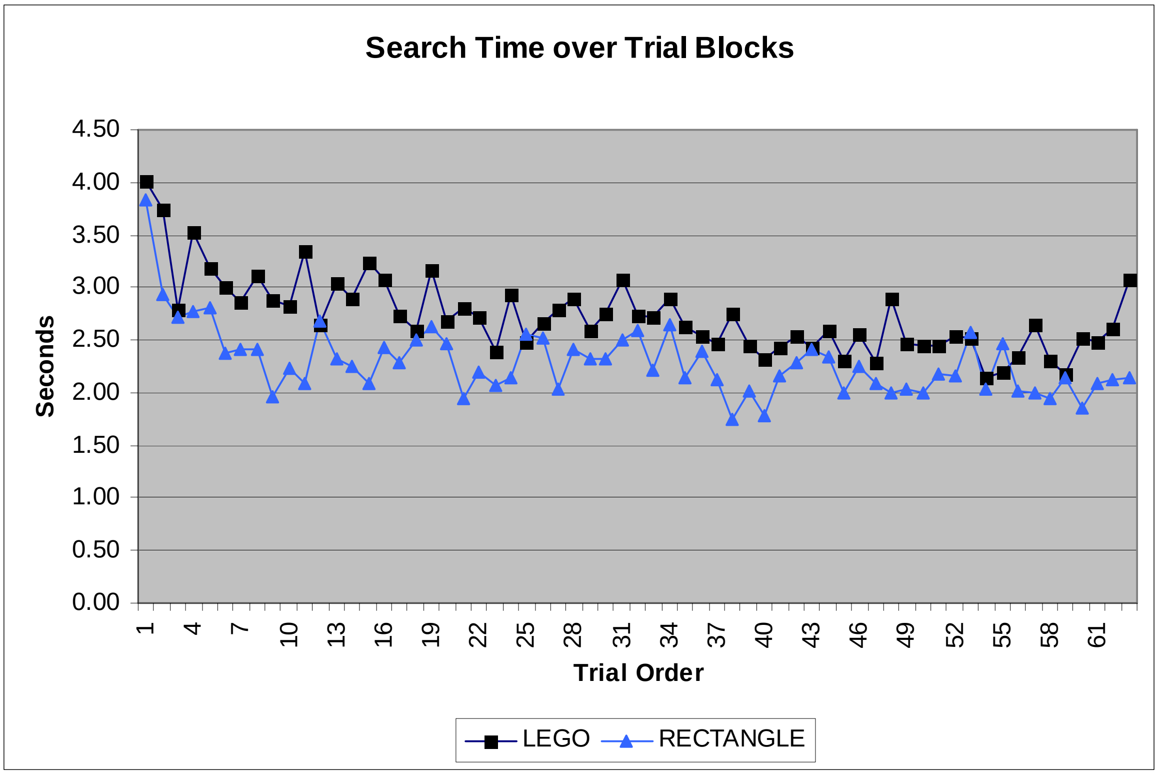



**Figure 8: Effect of time on task**

For both conditions, there appears to be an initial learning effect that stabilizes by about trial #8. Further, there appears to be an "end effect" such that performance in the Lego condition deteriorates on the last four trials. All of the major analyses reported below were re-run without the first seven trials, but this did not impact the pattern of results substantively. Therefore, these analyses have not been included in the report.

Given that the overall results demonstrated that the Rectangle three criteria search was the fastest, suggesting that it possibly formed a unitary icon or Gestalt, the question arises as to how quickly this may have been learned and is there any evidence to suggest similar learning for the Lego condition. The following graph shows search time over trials for the 3 criterion condition only for both designs. Since not all subjects did the same three criterion search at the same point in the session, the "n" on which mean is based varies considerably. To reduce some of the potential effects of this variation in the plotted data, means based upon one or two data points only have been omitted.

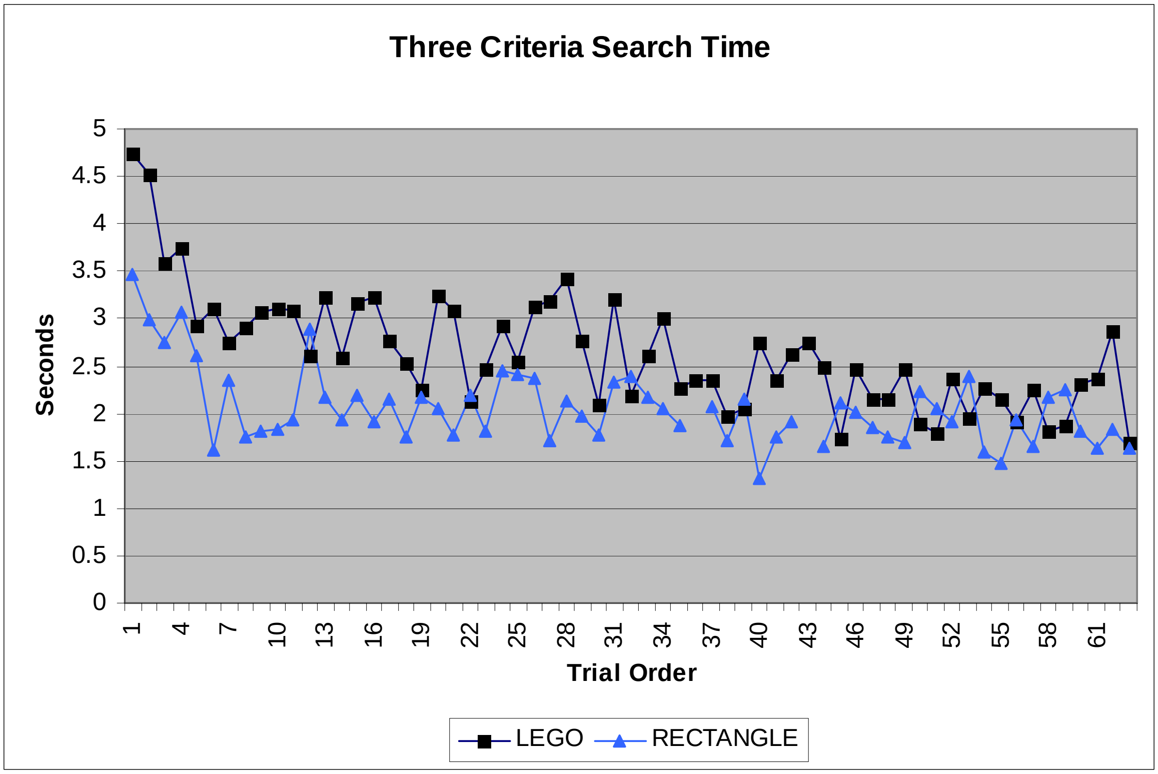



**Figure 9: Three criteria search time over trial blocks**

These data apparently show that the multiple properties of this configuration of the Rectangle icon are learned quite quickly and that search time remains relatively flat for the balance of the session. For the Lego condition, this learning effect appeared to take more trials, and were it not for three data points close to the end of the trial, there is evidence that the speed of search was approaching that of the Rectangle condition.

### 5.1.3 Combined Effects of Uncertainty Dimension, Level of Uncertainty and Icon Type

A more detailed analysis of the data was conducted to examine differences among conditions following the discovery of a significant 3 way interaction (icon type, level of uncertainty and number of search criteria) on search time ($F(4,48)=3.14$, $p=.02$).

First, we looked at performance differences across the levels of uncertainty taking into account factors of the uncertainty dimension and the icon type. Significantly different means are summarized in a table for each analysis and highlighted yellow. The critically significant difference required was .378 sec ($t(40)$, $p=.05$).

The following figures show mean search time for the two icon conditions for the different levels of uncertainty as a function of the number of search criteria.

**Time Uncertainty**

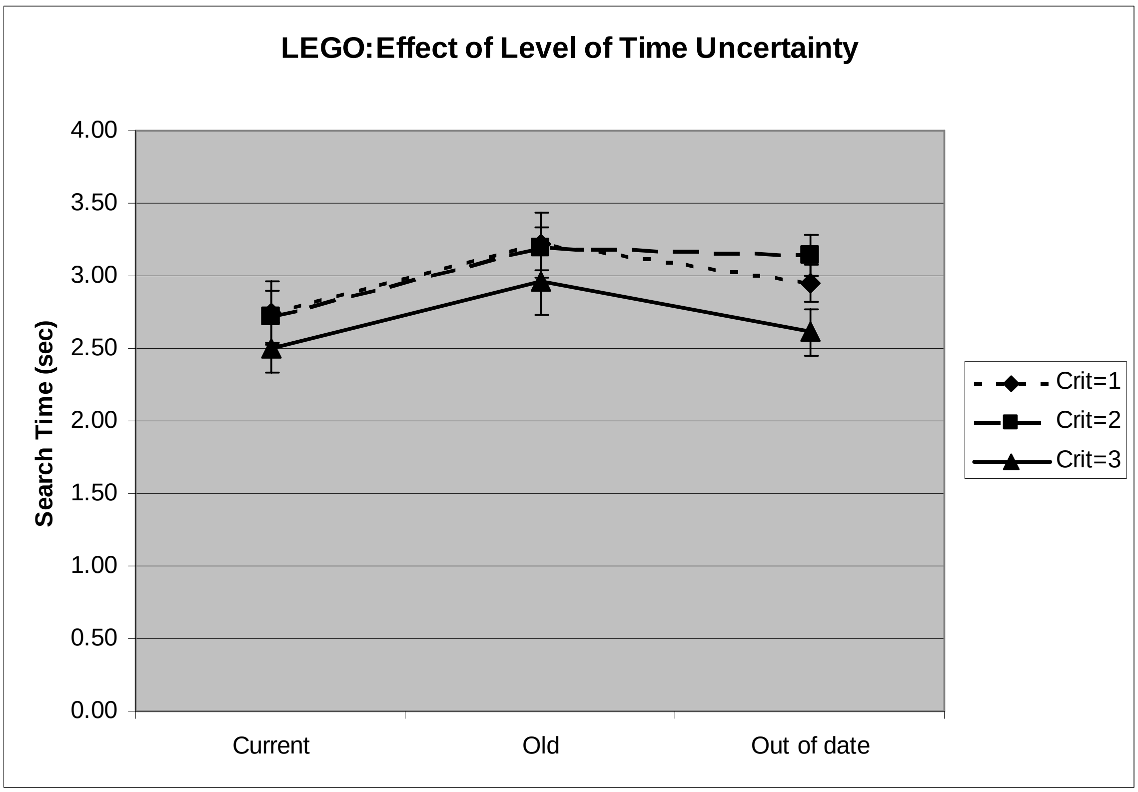



**Figure 10: LEGO: Search times for different time uncertainty icons**

**Table 7: Lego time uncertainty mean comparisons**

| Difference Pair | Old – Current | Out of Date – Current | Old – Out of date |
|---|---|---|---|
| One criterion | 0.47 | . | |
| Two criteria | 0.47 | 0.42 | |
| Three criteria | 0.46 | | |

**Summary:** The two block Lego condition produces slower search times than the single block condition for all levels of criterion search. The three block Lego condition also produces slower search times than the one block condition in the two criterion search condition.

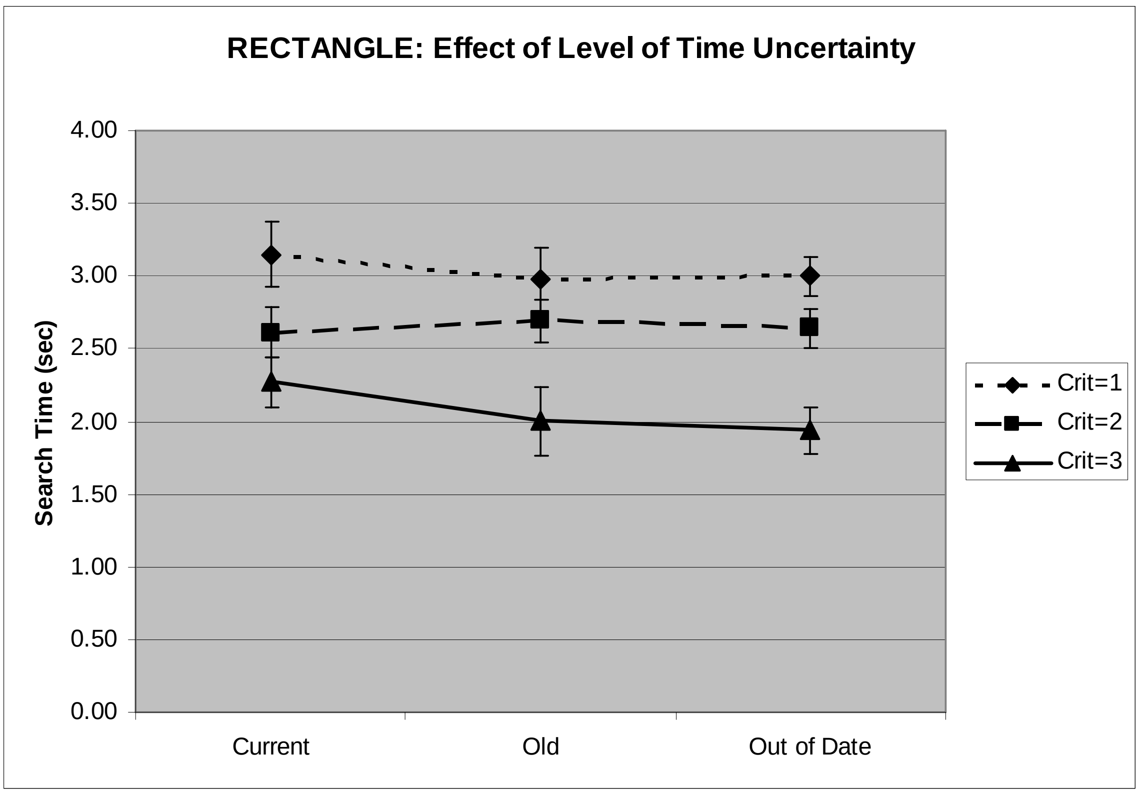



**Figure 11: RECTANGLE: Search times for different time uncertainty icons**

No significant differences were obtained between uncertainty levels for any of the number of search criteria conditions.

## Spatial Uncertainty

Similar data for the two icon conditions are shown below.

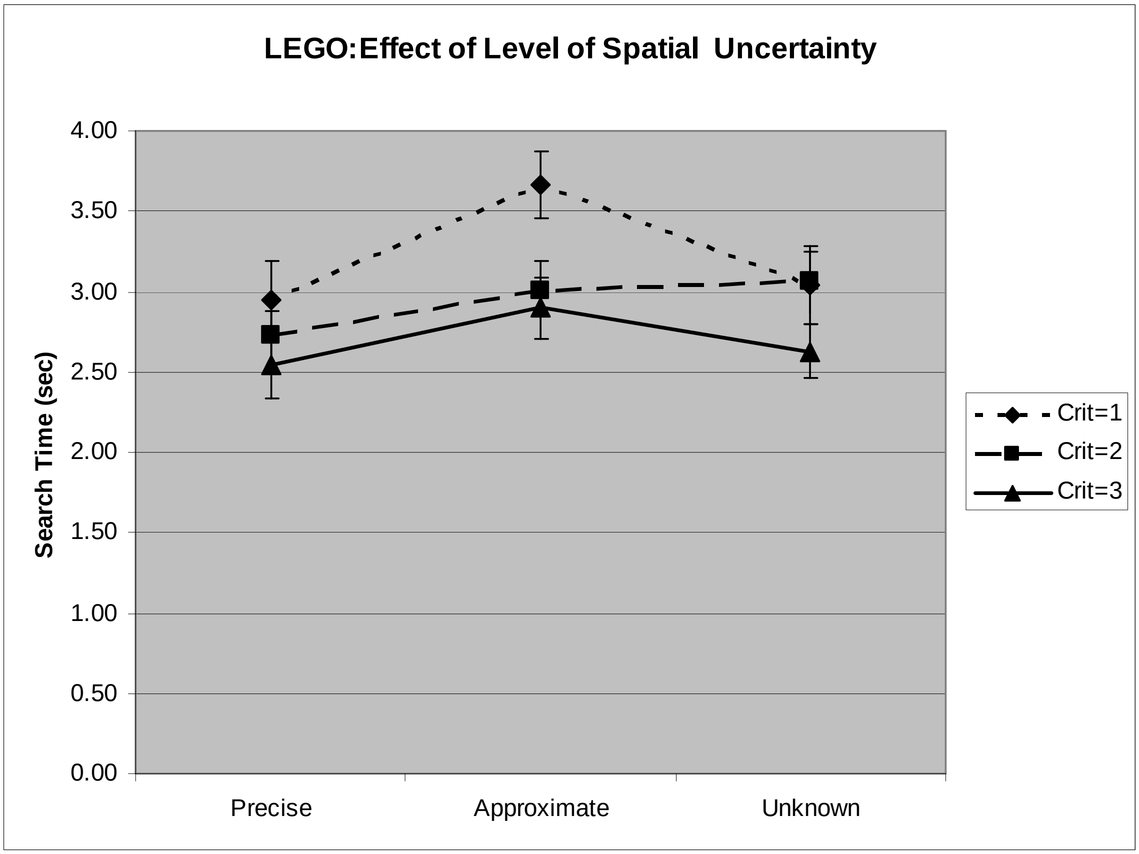



**Figure 12: LEGO: Search times for different spatial uncertainty icons**

**Table 8: Lego spatial uncertainty mean comparisons**

| Difference Pair | Approximate – Precise | Approximate – Unknown | Precise – Unknown |
|---|---|---|---|
| One criterion | .72 | .63 | |
| Two criteria | | | |
| Three criteria | | | |

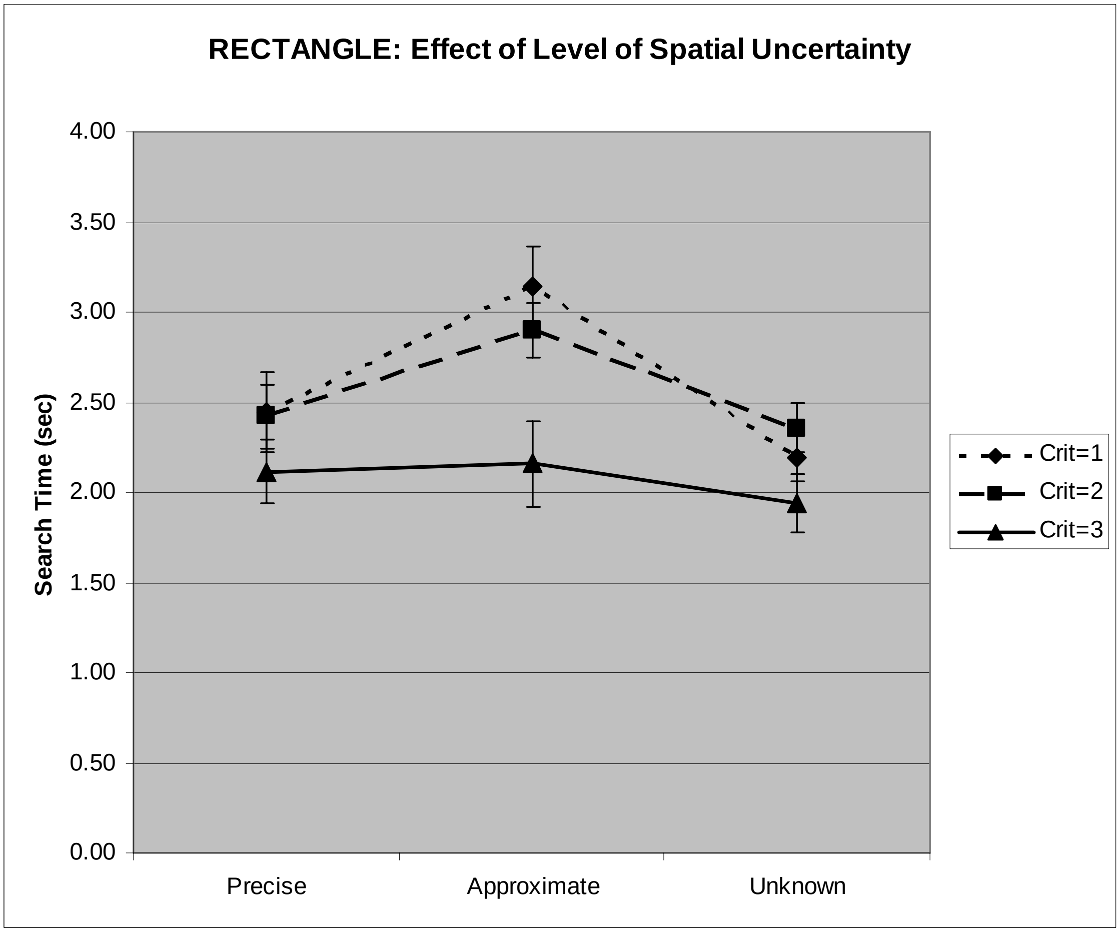



**Figure 13: RECTANGLE: Search times for different spatial uncertainty icons**

**Table 9: Rectangle spatial uncertainty mean comparisons**

| Difference Pair | Approximate – Precise | Approximate – Unknown | Precise – Unknown |
|---|---|---|---|
| One criterion | .70 | .94 | |
| Two criteria | .48 | .54 | |
| Three criteria | | | |

**Summary:** Again, the two block Lego produces slower search times than single block for a single criterion search. This is similar to the pattern of results for the time uncertainty (note that the block designs are equivalent except for a 90 degree rotation). In addition, the two block Lego condition also produced slower search times than the three block on this dimension. For the Rectangle design, differences are found both with single and two criteria searches. In this case, the intermediate sized circle produces slower response times than either the small or large circle in representing spatial uncertainty.

### Identity Uncertainty (colour differences)

Similar data for the two icon conditions are shown below.

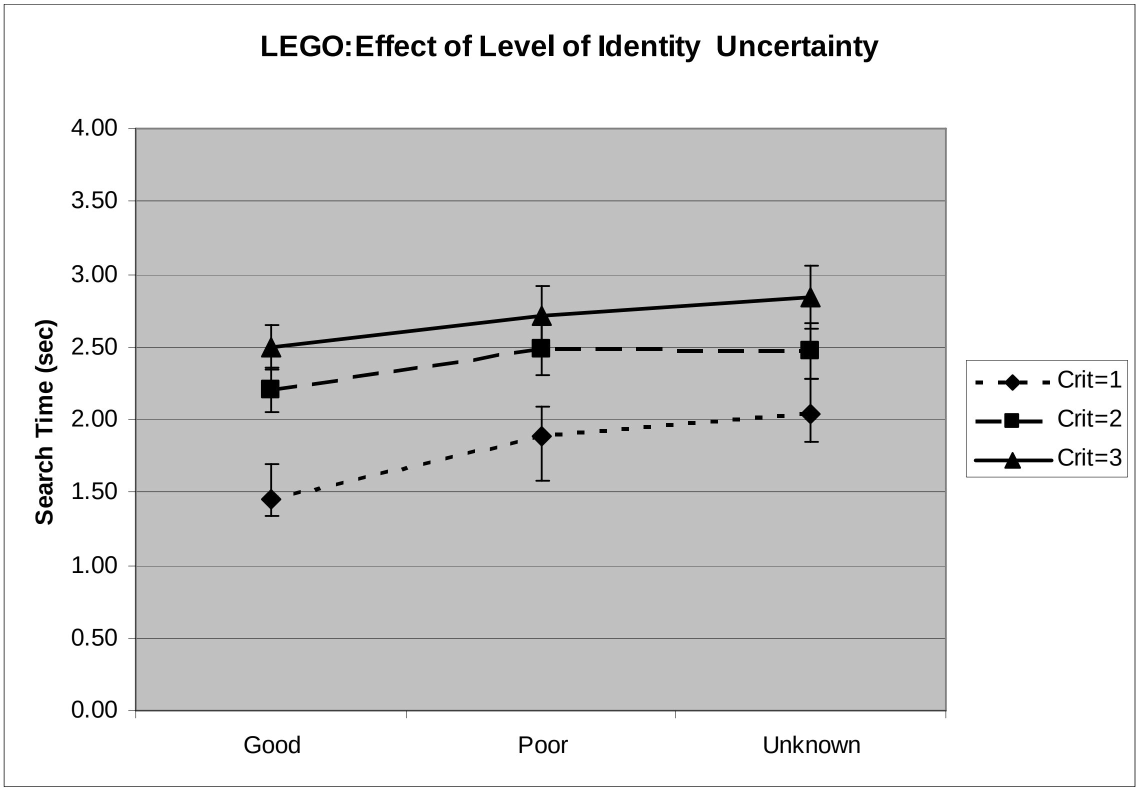



**Figure 14: LEGO: Search times for different identity uncertainty icons**

**Table 10: Lego identity uncertainty mean comparisons**

| Difference Pair | Poor – Good | Poor – Unknown | Unknown – Good |
|---|---|---|---|
| One criterion | .43 | | .59 |
| Two criteria | | | |
| Three criteria | | | |

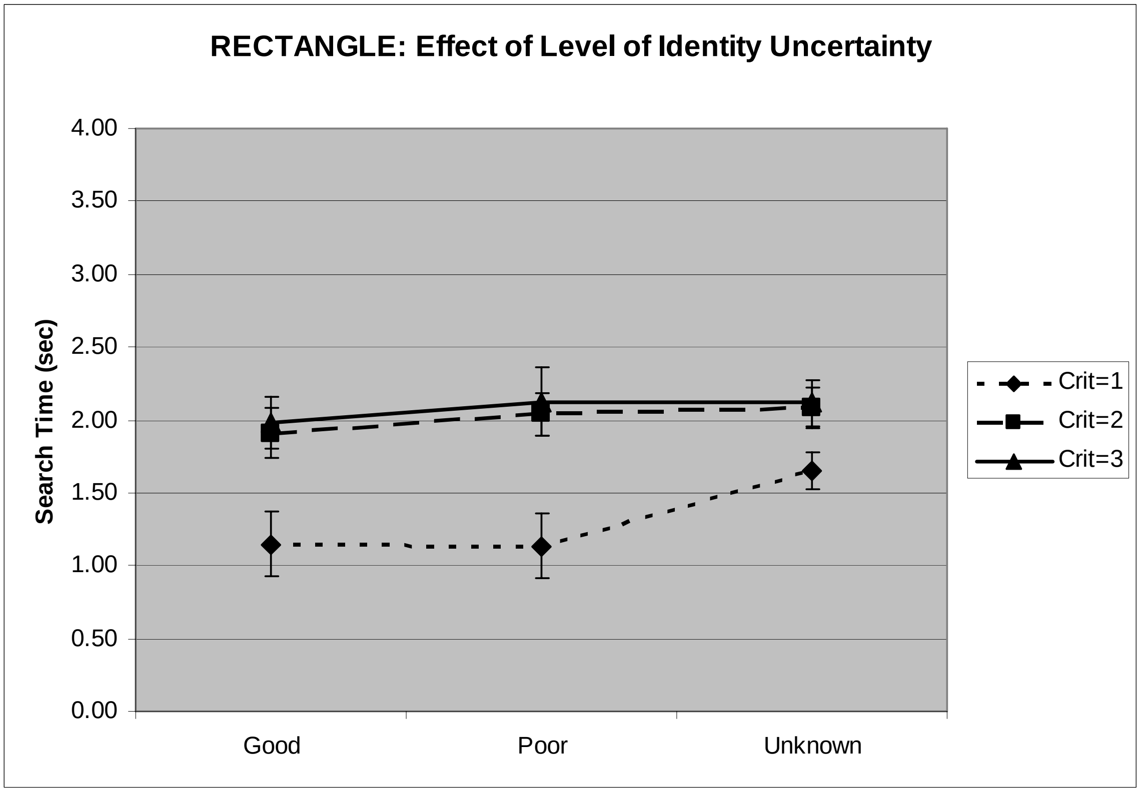



**Figure 15: RECTANGLE: Search times for different spatial uncertainty icons**

**Table 11: Rectangle identity uncertainty mean comparisons**

| Difference Pair | Poor – Good | Unknown – Poor | Unknown – Good |
|---|---|---|---|
| One criterion | | .52 | .51 |
| Two criteria | | | |
| Three criteria | | | |

**Summary:** For the Lego condition, the only search time differences were for the single criterion search, where the green icon was located more quickly than either the orange or yellow icons. In the Rectangle condition, again the only difference was with the single search criterion, in that the plum (unknown) colored background produced slower search times than the green or gold

Next we look at **differences between uncertainty dimensions**, taking into account the number of search criteria and the level of uncertainty. This analysis was done separately for the two icon designs. Note that instead of presenting the data redrawn again into the appropriate figures, we have extracted the relevant means for comparison and presented them in tabular format. Once again, yellow highlighted cells indicate a significant difference between the two means.

**Table 12: Lego Icon: comparison of uncertainty dimensions**

| | Single criterion | | | Two Criteria | | | Three Criteria | | |
|---|---|---|---|---|---|---|---|---|---|
| Uncertainty Level | Low | Medium | High | Low | Medium | High | Low | Medium | High |
| SPATIAL | 2.95 | 3.67 | 3.04 | 2.73 | 3.01 | 3.06 | 2.54 | 2.90 | 2.63 |
| TIME | 2.74 | 3.21 | 2.95 | 2.72 | 3.19 | 3.14 | 2.50 | 2.96 | 2.61 |
| ID | 1.45 | 1.88 | 2.03 | 2.20 | 2.48 | 2.47 | 2.50 | 2.72 | 2.84 |
| Spatial – Time | 0.21 | 0.45 | 0.09 | 0.01 | -0.18 | -0.08 | 0.03 | -0.07 | 0.01 |
| Spatial – ID | 1.50 | 1.78 | 1.00 | 0.52 | 0.53 | 0.59 | 0.04 | 0.18 | -0.21 |
| Time – ID | 1.30 | 1.33 | 0.92 | .51 | .71 | .67 | 0 | .24 | -0.23 |

As might be expected, with the exception of the single criterion/medium uncertainty condition, there were no differences between the search times for the vertical (time uncertainty) or horizontal (spatial uncertainty) elements of the design. In all other cases, except for the three criteria condition, identity search (colour coded) was always faster than searching for time or spatial uncertainty.

Similar comparisons for the Rectangle design are shown in the next table.

**Table 13: Rectangle Icon: comparison of uncertainty dimensions**

| | Single criterion | | | Two Criteria | | | Three Criteria | | |
|---|---|---|---|---|---|---|---|---|---|
| Uncertainty Level | Low | Medium | High | Low | Medium | High | Low | Medium | High |
| SPATIAL | 2.44 | 3.14 | 2.20 | 2.42 | 2.90 | 2.36 | 2.11 | 2.16 | 1.94 |
| TIME | 3.15 | 2.98 | 3.00 | 2.61 | 2.69 | 2.64 | 2.27 | 2.00 | 1.94 |
| ID | 1.15 | 1.13 | 1.65 | 1.91 | 2.04 | 2.08 | 1.98 | 2.12 | 2.12 |
| Spatial – Time | -0.71 | 0.16 | -0.80 | -0.19 | 0.20 | -0.28 | -0.16 | 0.16 | 0.00 |
| Spatial – ID | 1.29 | 2.01 | 0.54 | 0.51 | 0.86 | 0.27 | 0.14 | 0.04 | -0.18 |
| Time – ID | 2.00 | 1.85 | 1.34 | 0.70 | 0.65 | 0.56 | 0.29 | -0.12 | -0.18 |

The colour coded identity search was faster than searching for the spatial uncertainty indicator (circle size) or the time uncertainty indicator (circle position) for both single and two criterion search conditions (with the exception of the "high" uncertainty condition for the latter). Of the spatial versus time uncertainty comparisons, the only differences occurred for the single criterion search where searching for circle size was faster than searching for circle location, when the latter was at either end of the background rectangle. However, the meaningfulness of comparing search times for a small circle with the location of that circle could be questionable.

Of note, from the above analyses is the observation that the differences in search time performance between the various components of the uncertainty design icons appear to wash out under the three criteria search condition. Again, this might suggest that subjects were able to encode the icon more holistically as a unified gestalt.

The final exploration of the three way interaction examines differences between the icon designs taking into account the factors of uncertainty level and the number of search criteria. The relevant means are shown in the next three tables, for each of the uncertainty dimensions separately, starting with spatial uncertainty.

**Table 14: Spatial uncertainty search times for each icon design**

| | Single criterion | | | Two Criteria | | | Three Criteria | | |
|---|---|---|---|---|---|---|---|---|---|
| Uncertainty Level | Precise | Approximate | Unknown | Precise | Approximate | Unknown | Precise | Approximate | Unknown |
| LEGO | 2.95 | 3.67 | 3.04 | 2.73 | 3.01 | 3.06 | 2.54 | 2.90 | 2.63 |
| RECTANGLE | 2.44 | 3.14 | 2.20 | 2.42 | 2.90 | 2.36 | 2.11 | 2.16 | 1.94 |
| Difference | 0.51 | 0.53 | 0.84 | 0.31 | 0.11 | 0.70 | 0.43 | 0.74 | 0.69 |

With the exception of two comparisons in the two criteria search conditions, the Rectangle design always resulted in significantly faster search times than the Lego design, suggesting that the circle size representation of uncertainty is a better choice than the horizontal bar indicator in the Lego design.

Similar data for time uncertainty search are shown in the next table.

**Table 15: Time uncertainty search times for each icon design**

| | Single criterion | | | Two Criteria | | | Three Criteria | | |
|---|---|---|---|---|---|---|---|---|---|
| Uncertainty Level | Current | Old | Out of date | Current | Old | Out of date | Current | Old | Out of date |
| LEGO | 2.74 | 3.21 | 2.95 | 2.72 | 3.19 | 3.14 | 2.50 | 2.96 | 2.61 |
| RECTANGLE | 3.15 | 2.98 | 3.00 | 2.61 | 2.69 | 2.64 | 2.27 | 2.00 | 1.94 |
| Difference | -0.41 | 0.23 | -0.05 | 0.11 | 0.50 | 0.50 | 0.23 | 0.96 | 0.67 |

Here the results are less consistent. Depending upon the number of criteria, search for the location of the circle may or may not be faster than searching for the number of squares in the vertical bar of the Lego design.

Finally, we look at differences in identity search times, which for both designs was indicated by colour coding.

**Table 16: Identity uncertainty search times for each icon design**

| | Single criterion | | | Two Criteria | | | Three Criteria | | |
|---|---|---|---|---|---|---|---|---|---|
| Uncertainty Level | Good | Poor | Unknown | Good | Poor | Unknown | Good | Poor | Unknown |
| LEGO | 1.45 | 1.88 | 2.03 | 2.20 | 2.48 | 2.47 | 2.50 | 2.72 | 2.84 |
| RECTANGLE | 1.15 | 1.13 | 1.65 | 1.91 | 2.04 | 2.08 | 1.98 | 2.12 | 2.12 |
| Difference | 0.30 | 0.75 | 0.38 | 0.29 | 0.44 | 0.39 | 0.52 | 0.60 | 0.72 |

With the exception of two conditions, where the trend was similar but not statistically significant, it appears that the color coding scheme used for the Rectangle design produces consistently faster identity searches than the color scheme for the Rectangle design. This may have been due to the specific colors that were used, or to the fact that the size of the color area for the Rectangle design was larger than that of the Lego design.

## 5.2 Accuracy

Accuracy data are shown in the following figures. In general accuracy was very high and ANOVA failed to show any significant differences among conditions. The lower accuracy in the single criterion condition is somewhat misleading as it was caused by higher error rates for two subjects (note that nine of the thirteen subjects scored 100% and one 97%). The data have been re-plotted in the next figure omitting these two subjects.

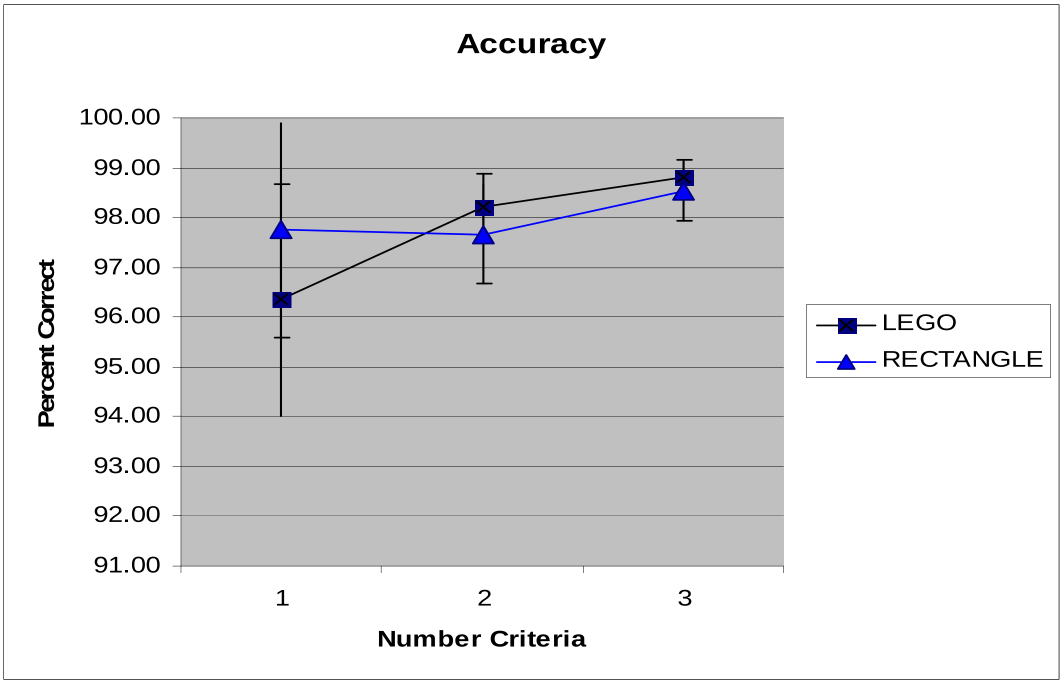



**Figure 16: Percent correct identifications**

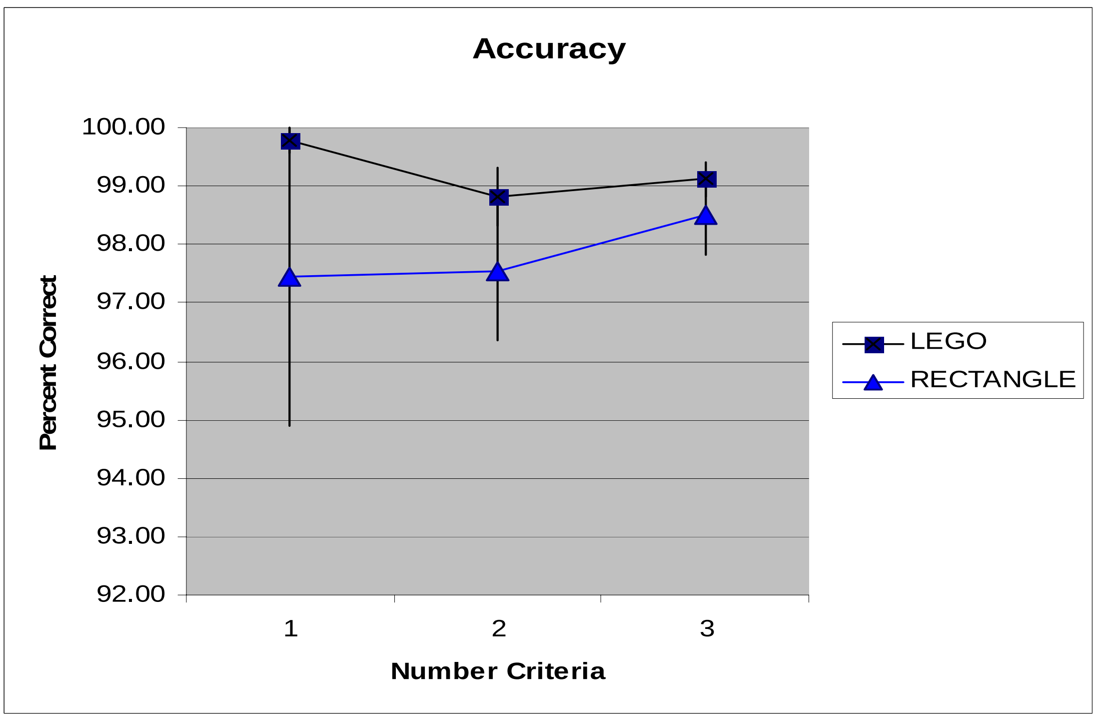



**Figure 17: Percent correct identifications: two outliers removed**

## 5.3 "Think Time"

We have defined the time between the subjects receiving the instruction criteria and the point at which they start the trial as "think time", i.e. they are constructing a mental image of the characteristics of the target for which they will be searching. If one design presents greater difficulty in this respect, then we should expect to see a difference in think time between the conditions.

The relevant data are shown in the next figure.

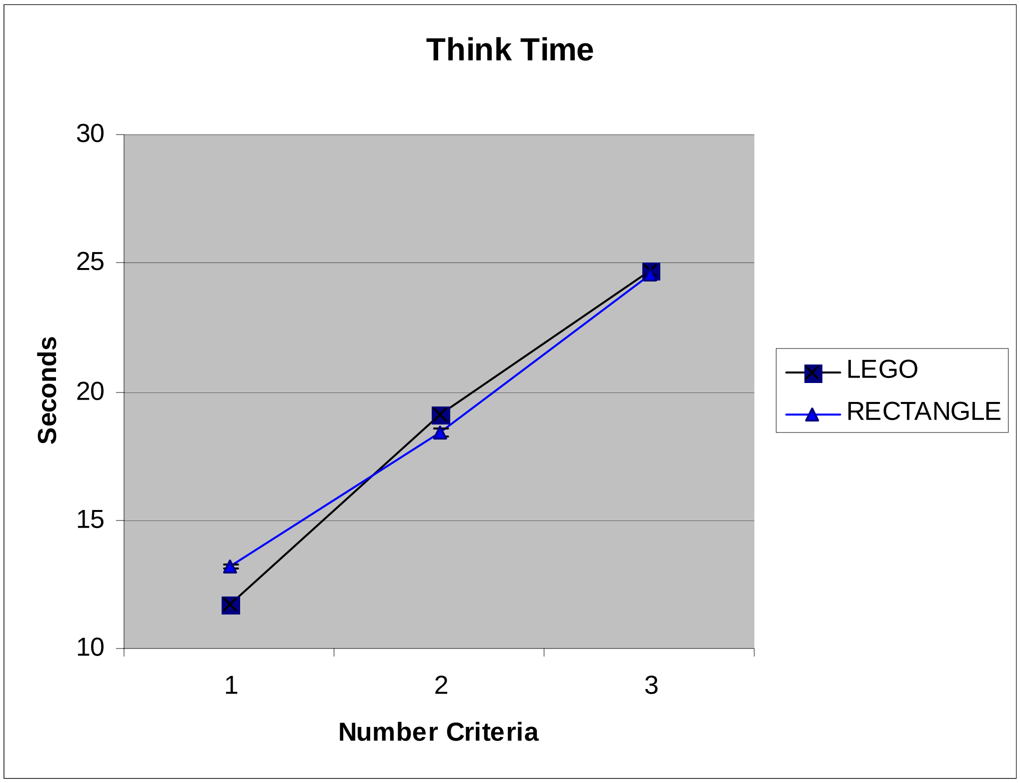



**Figure 18: Mean think time**

Clearly, both designs represent the same level of difficulty, and the function shows a clear linear increase with each additional aspect of uncertainty that must be considered. This was supported by the ANOVA showing only a main effect for the number of search criteria ($F(2,24)=42.7$, $p<.001$). In spite of the similarity of the group means, inspection of the individual subject data in the following graph shows that there is significant individual variation from the group trend. The data below are shown as the time difference in think time between the Rectangle and Lego conditions. Positive values indicate that the Rectangle condition was slower, negative values that the Lego condition was slower. Five subjects showed consistently (i.e. all 3 criteria) slower think times for the Lego condition, whereas one subject showed consistently slower times for the Rectangle condition.

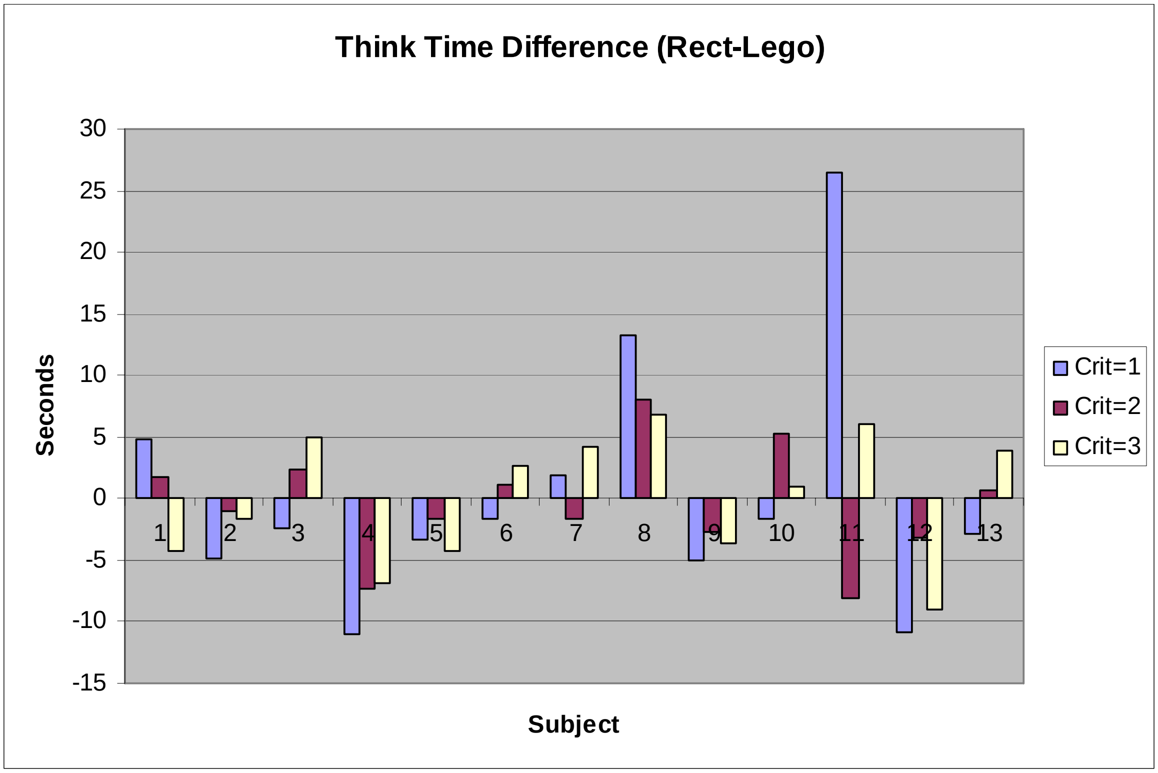



**Figure 19: Individual subjects: difference in think time between designs**

The next figure shows how think time changes over the course of the experiment and demonstrates a clear effect of learning of the specific icon patterns. There appears to be no discernable evidence from these data that one design is learned more quickly than the other. The flattening of the curve suggests that subjects had learned the pattern of the icons towards the latter part of the test session.

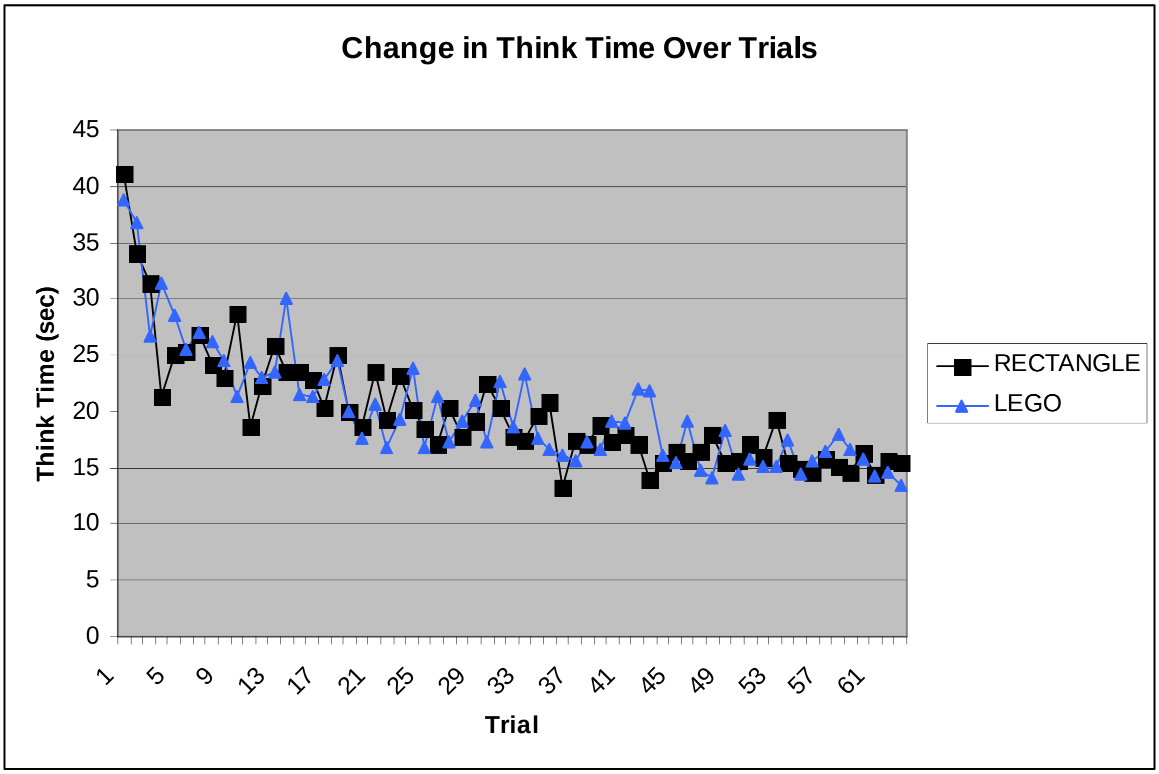



**Figure 20: Think time as a function of time on task**

## 5.4 Stop Time

We have called the time between identifying the last target selected on a trial and the subject indicating that they are done the “stop time”. This presumably represents a period of search when the subject is scanning the display to determine if all targets have been located. The following figure shows the mean stop time per trial for both design conditions. Stop time was almost significantly faster for the Rectangle condition ($F(1,12)=4.61$, $p=.052$) than the Lego condition and was faster with increasing number of search criteria for both conditions ($F2,24)=9.58$, $p<.001$). This is a welcome result as it suggests that subjects are able to use “gestalt” properties of the icon when searching for multiple criteria, rather than looking individually for the separately coded uncertainty dimensions (which would presumably have shown increased times with increasing number of criteria).

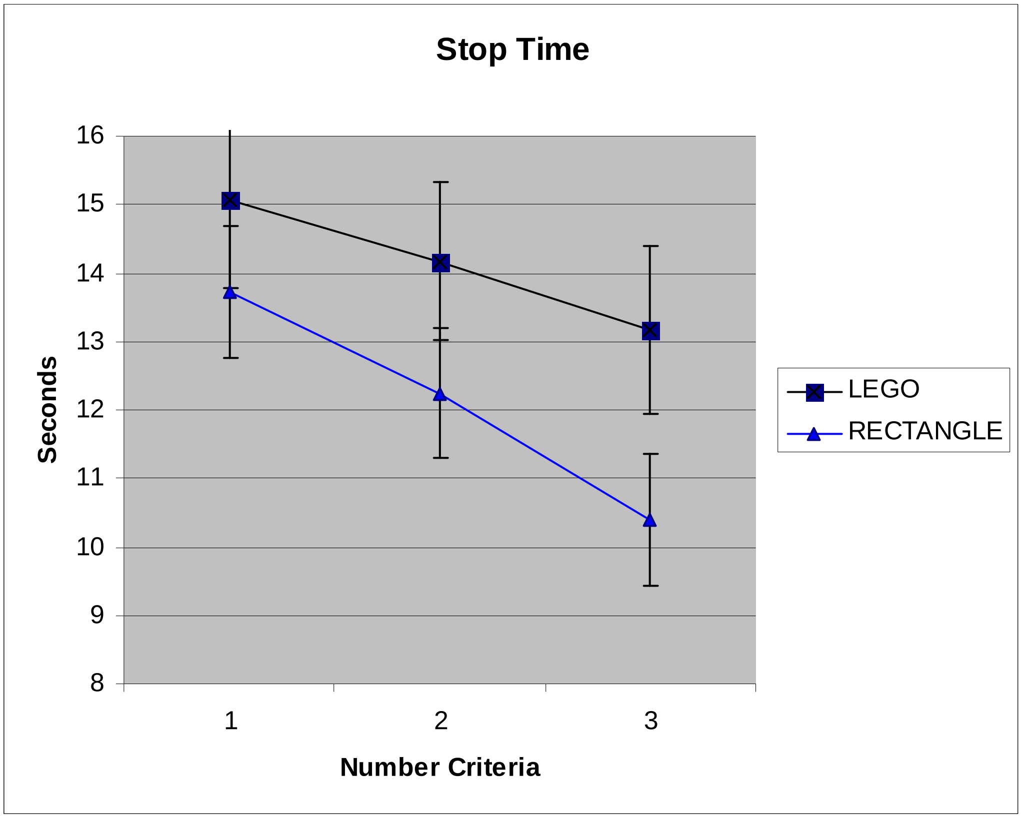



**Figure 21: Mean stop time**

## 5.5 Questionnaire Data

Subjects performed questionnaire ratings at the end of each session. These are shown in the following tables. For questions **1-8**, the following instructions were provided:

Rate your **agreement** with the following statements, using the following scale:

1. Completely disagree
2. Somewhat disagree
3. Neither agree or disagree
4. Somewhat agree
5. Completely agree

Questions **9 and 10** dealt with workload using the following scale.

Use the following scale to rate the **workload** you experienced for the areas below:

1. The load was very low and I could do this task continuously with little fatigue
2. The load was moderately low and I could do this task for a few hours at a time
3. The load was noticeable and I would need regular breaks from the task

4. The load was difficult but I could handle it for the duration of the experiment
5. The load was very high and there were times when I would have liked to stop.

**Table 17: Ratings for task difficulty and workload**

| | | Lego | | Rectangle | |
|---|---|---|---|---|---|
| | Statement | Mean Rating | St. Dev | Mean Rating | St. Dev |
| 1 | The different levels of identity information were easy to comprehend. | 3.92 | 1.38 | 4.62 | 0.51 |
| 2 | The different levels of the spatial information were easy to comprehend. | 4.08 | 1.26 | 4.23 | 1.17 |
| 3 | The different levels of the time information were easy to comprehend. | 4.23 | 0.93 | 4.31 | 0.95 |
| 4 | The visual representations of the different levels of the identity information were easy to discriminate. | 4.38 | 1.12 | 4.77 | 0.44 |
| 5 | The visual representations of the different levels of the spatial position information were easy to discriminate. | 4.31 | 0.85 | 3.77 | 1.24 |
| 6 | The visual representations of the different levels of the time information were easy to discriminate. | 4.31 | 0.85 | 4.62 | 0.65 |
| 7 | It was easy to know what to look for when there were multiple criteria. | 4.54 | 0.52 | 4.54 | 0.52 |
| 8 | It was easy to find the contacts when there were multiple criteria. | 4.08 | 0.95 | 4.38 | 0.65 |
| 9 | The task of thinking about what to look for based on the search criteria. | 2.54 | 1.20 | 1.92 | 0.95 |
| 10 | The task of visually trying to find each contact. | 2.69 | 1.25 | 1.69 | 0.75 |

It should be noted that no statistical analysis has been performed on any of the questionnaire data at present. In general, the ratings for questions 1-8 showed a slightly more favourable assessment of the Rectangle design, with the noticeable exception of question 5, concerning the representation of spatial information. (Although inspection of the data showed that there were three subjects whose data were strongly in the reverse direction from the rest of the group and were responsible for this trend). Also workload ratings were lower for the Rectangle design.

Subjects also reported on any potential visual discomfort areas, using the following scale.

"Using the following scale, rate your experience with any of the following symptoms of **visual discomfort** during the course of the experiment "

1. Not at all
2. Occasionally, did not bother me
3. Sometimes and it was quite noticeable
4. Frequently and it bothered me
5. A lot and I felt significant discomfort

**Table 18: Reported frequency of visual discomfort**

| | Lego | | Rectangle | |
|---|---|---|---|---|
| Statement | Mean Rating | St. Dev | Mean Rating | St. Dev |
| Dry, itching or sore eye | 1.69 | 0.75 | 1.46 | 0.66 |
| Blinking | 1.46 | 0.78 | 1.31 | 0.48 |
| Watery eye | 1.15 | 0.38 | 1.08 | 0.28 |
| Pain around the eye area | 1.15 | 0.38 | 1.00 | 0.00 |
| Headache | 1.08 | 0.28 | 1.00 | 0.00 |
| Other muscle aches | 1.62 | 0.65 | 1.46 | 0.66 |

Again, the reporting pattern seems to favor the Rectangle design, although in both cases the absolute ratings are sufficiently low that they indicate that neither design would produce significant visual discomfort, at least over a duration of about 90 minutes of intensive activity.

# 6. Phase II Results: Target Search with Sensor Overlay Swaths

To review, this phase of the experiment involved search for either of the two design concepts against a criterion area (swath) that was characterized by its quality or time lateness. Two design options to represent quality were evaluated (a grey or hash fill), and two options for lateness (border solid/broken or border colour).

## 6.1 Target Stimuli

Only a subset of the complete 63 possible combinations of target stimuli were used for this phase (see Section 4.6) and these were randomly selected for each trial. The frequency of occurrence of these stimuli for each of the Lego and Rectangle conditions are shown in the following figure. The abbreviations for the target conditions are based on Table 4, with "A" indicating an absence or no criterion for an uncertainty dimension.

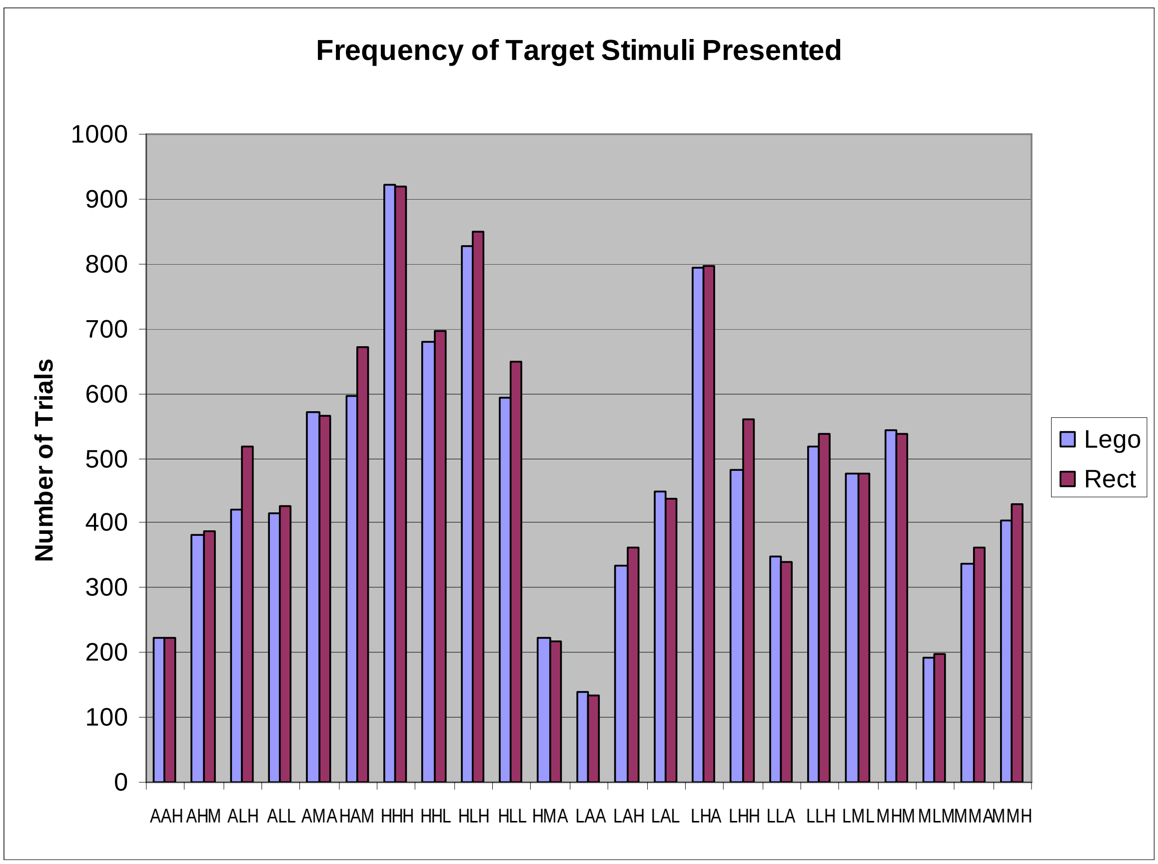



**Figure 22: Target stimuli used in phase II**

Although the actual frequencies with which the targets were selected varied considerably, the same pattern of selection was evident for both the Rectangle and Lego conditions, thereby making the results of comparisons across these conditions equivalent.

## 6.2 Target Search time

### 6.2.1 Number of criteria

Mean search time as a function of the number of criteria is shown in the next figure, where the pattern of results is very similar to phase I of the experiment. ANOVA showed a significant interaction between the icon design and the number of search criteria ($F(2,20)=8.66, p<.001$). Comparisons of individual means (critical difference =.15 sec, $p<.01$), showed that the only significant differences between the Lego and Rectangle conditions were for the three criterion search and that for the Rectangle condition the three criteria search was faster than the two criteria search.

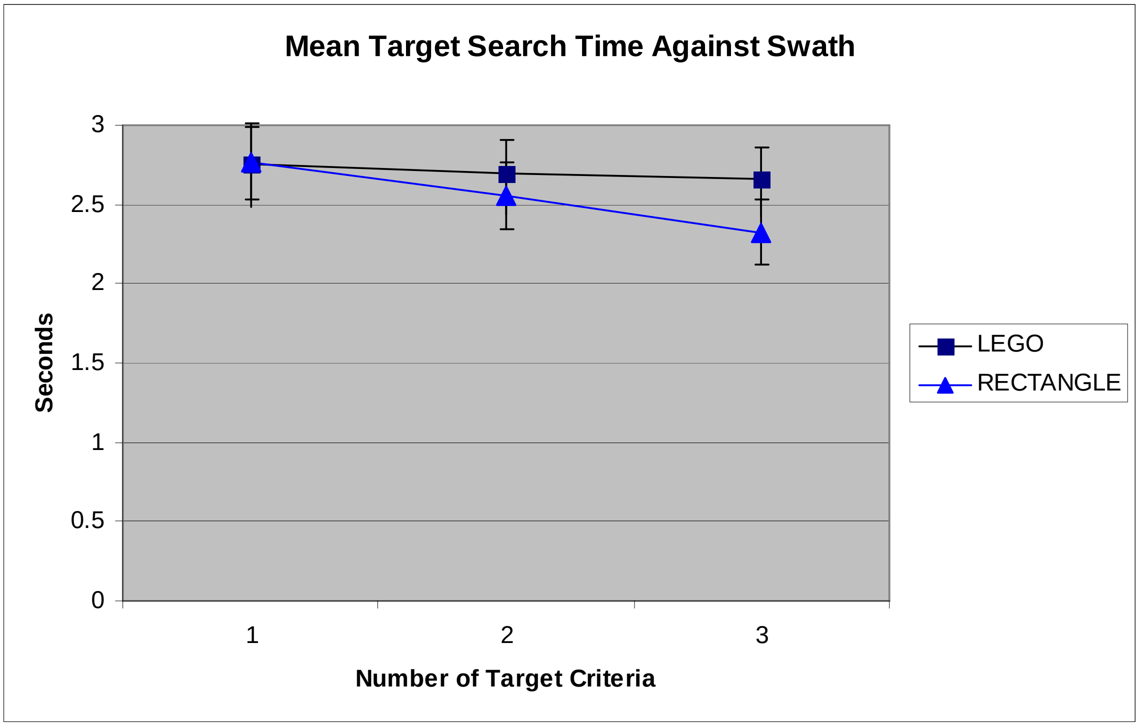



**Figure 23: Target search time as a function of design type and number of search criteria**

In comparing these results with Phase I of the experiment with no swath backgrounds, the mean search time per target for the Lego condition was virtually identical (2.70 versus 2.71 seconds in phase I). However, the Rectangle condition was slightly slower (2.55 versus 2.25 seconds in Phase I). We have no particular explanation for this result.

## 6.3 Effect of swath design

Search performance is shown in the following figures for each of the target design concepts against each of the swath configurations. The trials where the border was a green or gold color are indicated by “color” and those where the border was solid or broken are indicated by “line”. The trials when the background was either a light or darker shade of grey are indicated by “grey” and those where there was a cross hatch are indicated by “hash”. For the Lego condition, the performance is very similar across all backgrounds, but the two conditions with the grey background produce slightly faster search times than the hash background. Similarly, the two conditions with the line border produce faster times than with the colour border.

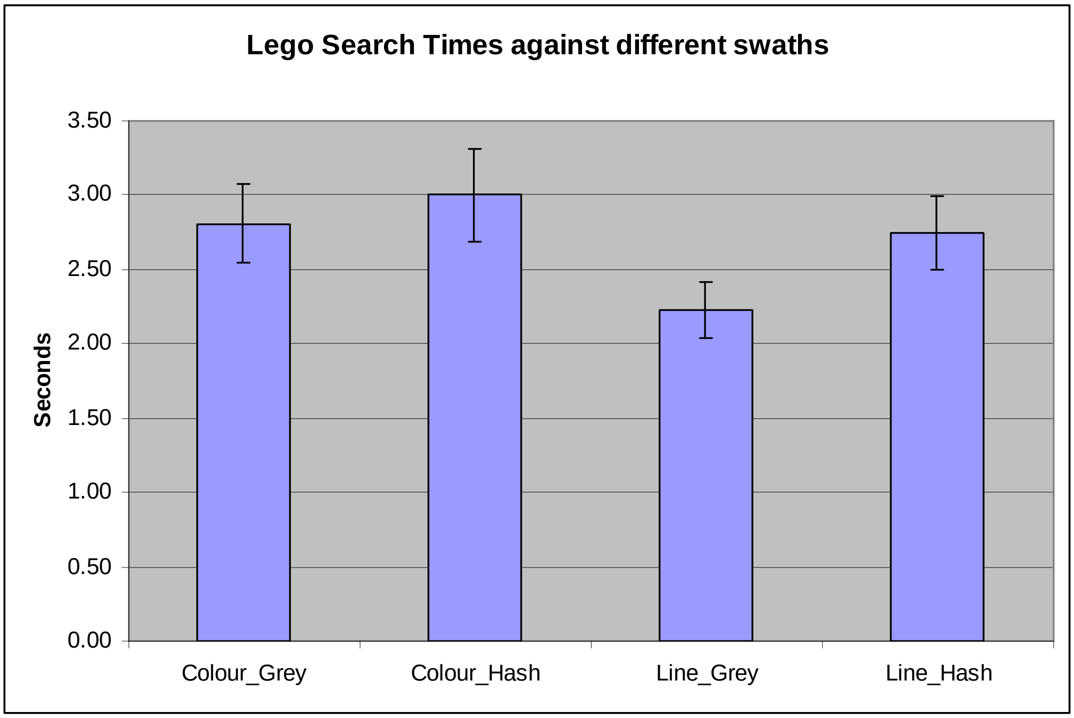



**Figure 24: Lego search times against different swaths**

A similar pattern is found for the Rectangle icon, as shown below.

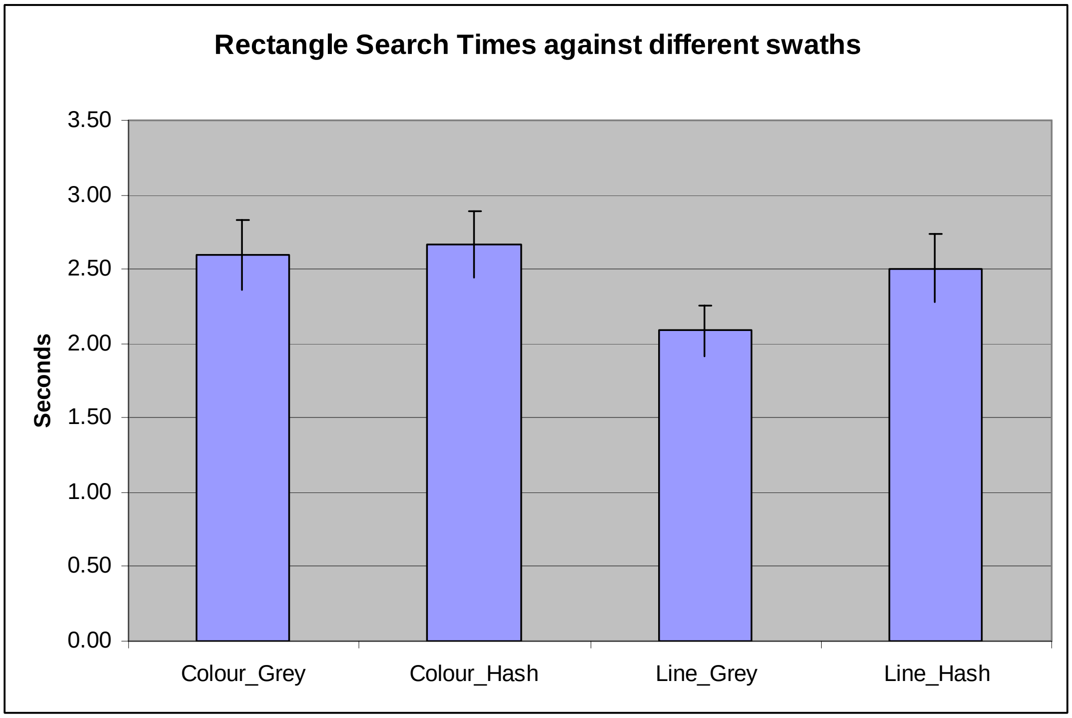



**Figure 25: Rectangle search time against different swaths**

ANOVA showed both a main effect for the type of sensor quality design (F(1,10)=25.05, p<.001) and sensor time uncertainty design (F(1,10)=27.58, p<.001), but no significant interaction between these variables. This means that for both icon types the grey background produced faster searches (2.43 versus 2.73 seconds) than the hashed background and the line border faster times than colored lines (2.38 versus 2.77 seconds).

### 6.3.1 Accuracy

Accuracy rates were very high for both the Lego and Rectangle targets for all of the sensor swath conditions, as shown in the following graphs. ANOVA revealed that there are no statistically significant differences in search performance between the target design conditions and the two design concepts for each of sensor quality and time lateness.

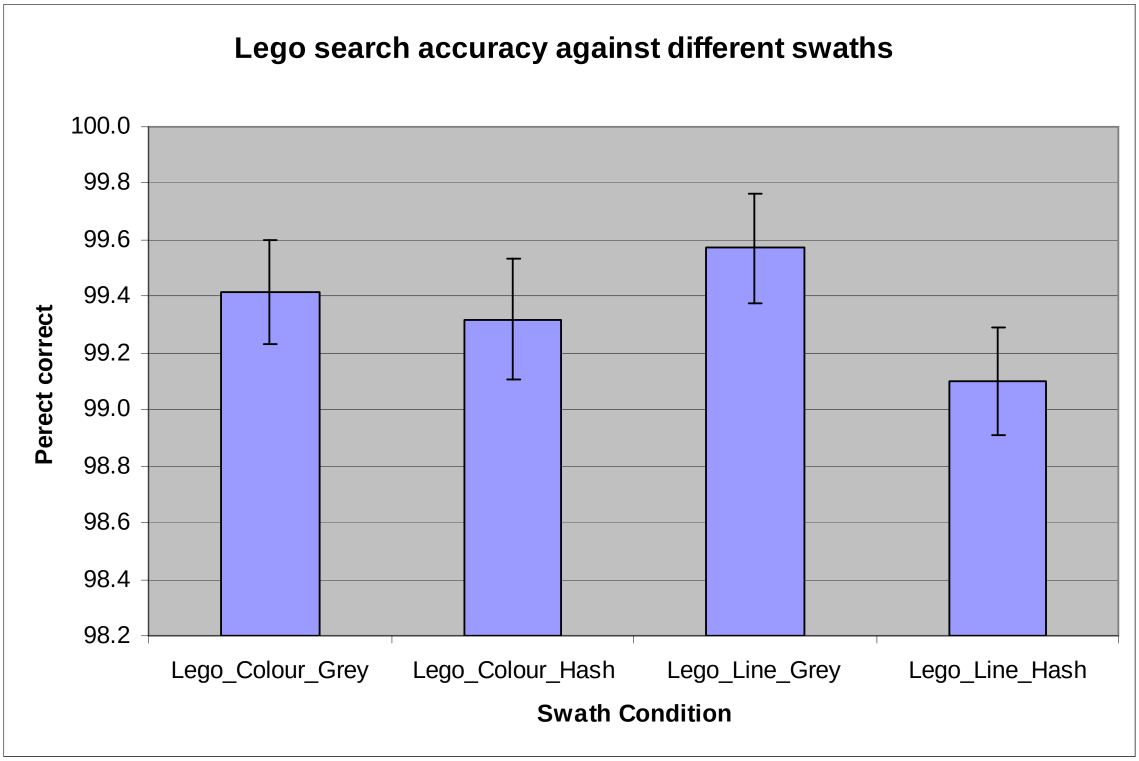



**Figure 26: Lego search accuracy against different swaths**

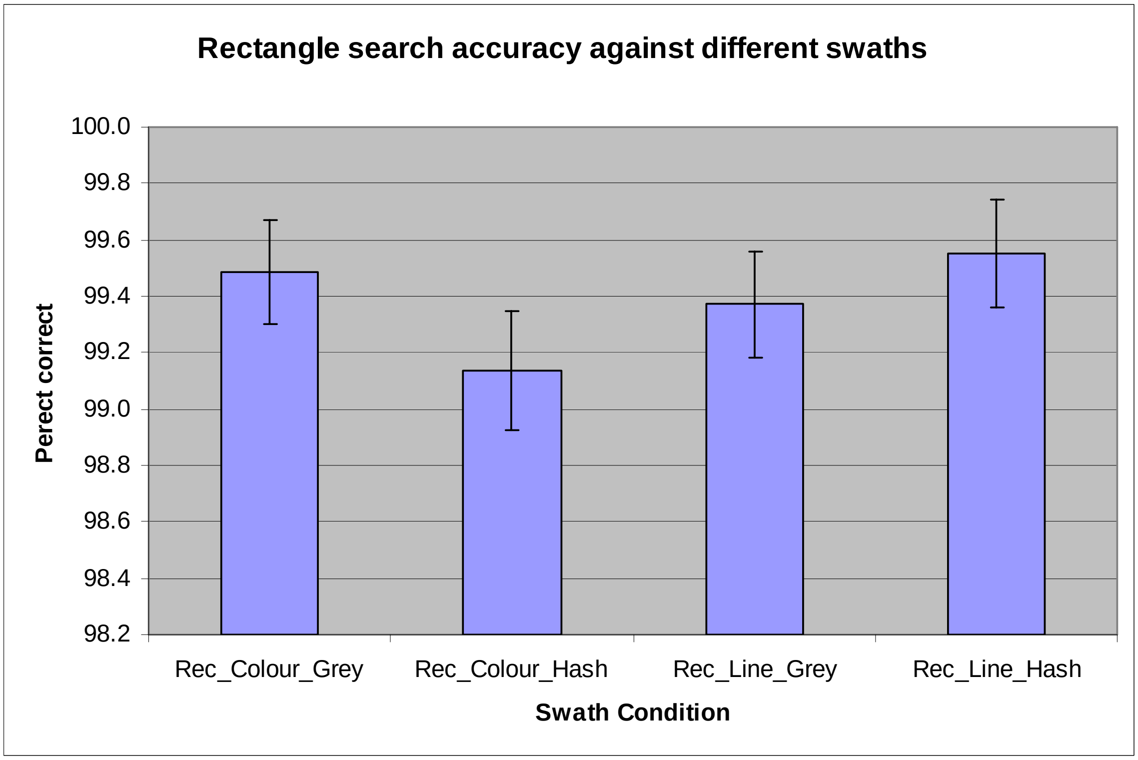



**Figure 27: Rectangle search accuracy against different swaths**

### 6.3.2 "Think" Time

As before, the "think" time is the latency between the subject being presented with the search criteria and then pressing the key to start the trial. Think time for the Rectangle and Lego conditions against the different swath combinations are shown separately in the next two figures.

As can be seen the differences among the swath conditions are very small and overall the Rectangle condition shows slightly shorter times than the Lego condition. Results from the ANOVA showed that indeed the difference between the Rectangle and Lego targets was statistically significant, $F(1,10)=5.04$, $p=.048$). In addition, think times for targets with the line method to indicate time uncertainty was statistically faster than with the colored border ($F(1,10)=7.59$, $p=.02$). There was no statistical difference in think time performance between the grey area fill and the cross hatching.

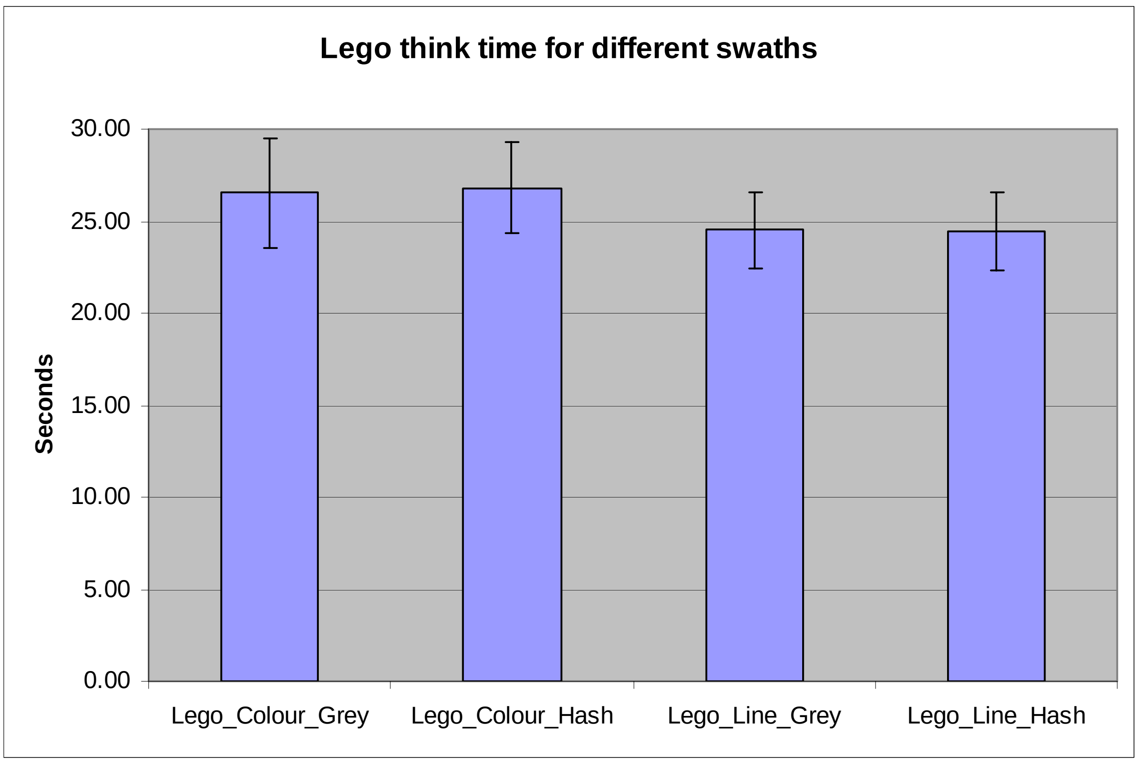



**Figure 28: Lego "think" time for different swaths**

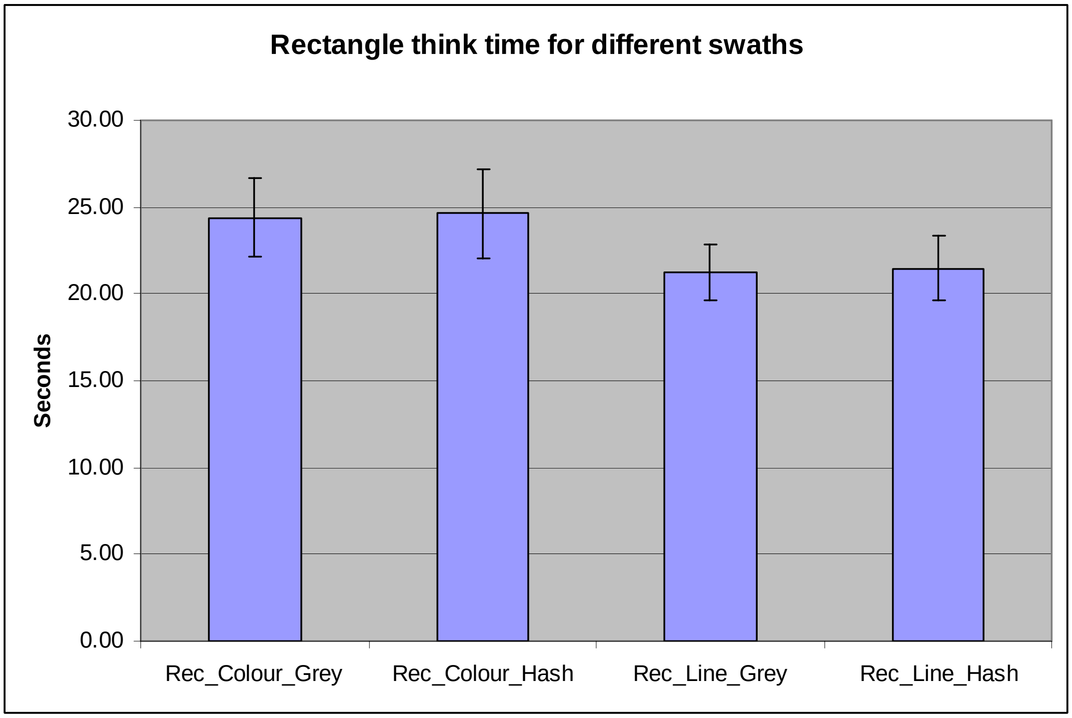



**Figure 29: Rectangle "think" time for different swaths**

The following table compares the mean think time in phases I and II.

**Table 19: Comparison of "think" time (seconds) with and without swath criteria**

| | **Lego** | **Rectangle** |
|---|---|---|
| Phase I (no swath) | 20.2 | 19.8 |
| Phase II (swaths) | 25.6 | 22.9 |

For the Lego condition the need to consider the additional background swath criteria adds about 27% to the thinking time, for the Rectangle condition the increase is approximately 16%.

### 6.3.3 Stop Time

Stop time is the latency between the time at which the last target is found and the subject indicating that the trial is over because they think they have found all of the targets. Stop time for the Rectangle and Lego conditions against the different swath combinations are shown separately in the next two figures. ANOVA showed again that the Rectangle icon was searched faster than the Lego design ($F(1,10)=9.16$, $p=.013$). Further, the line method for representing time lateness again resulted in faster times than the colored border ($F(1,10)=5.96$, $p=.03$) and the grey area fill also produced faster times than the cross hatching ($F(1,10)=15.63$, $p=.002$).

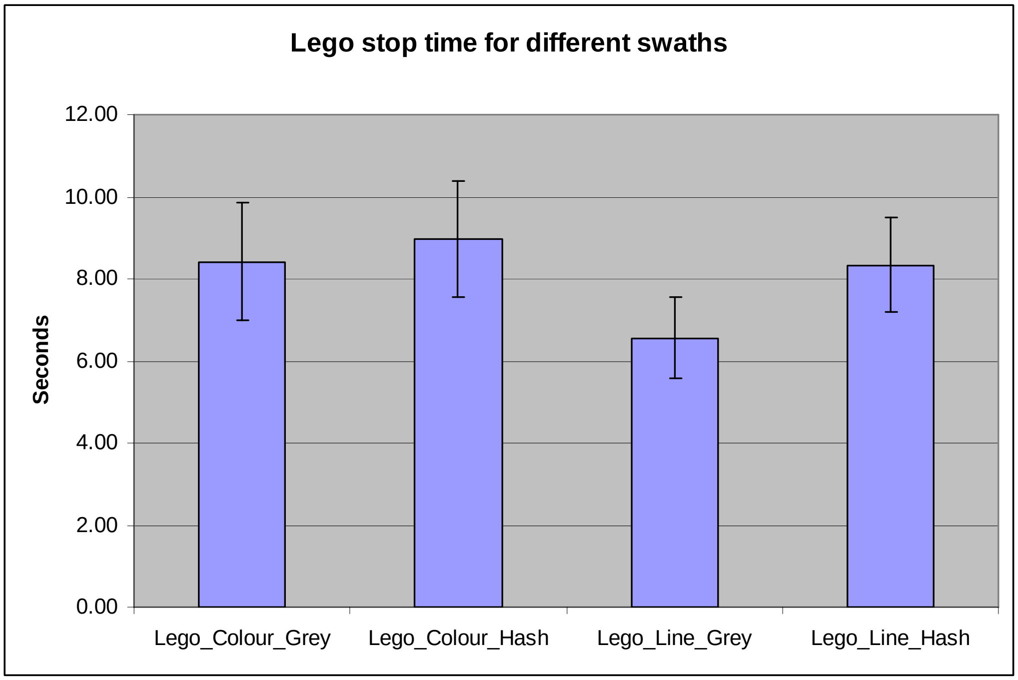



**Figure 30: Lego stop time for different swaths**

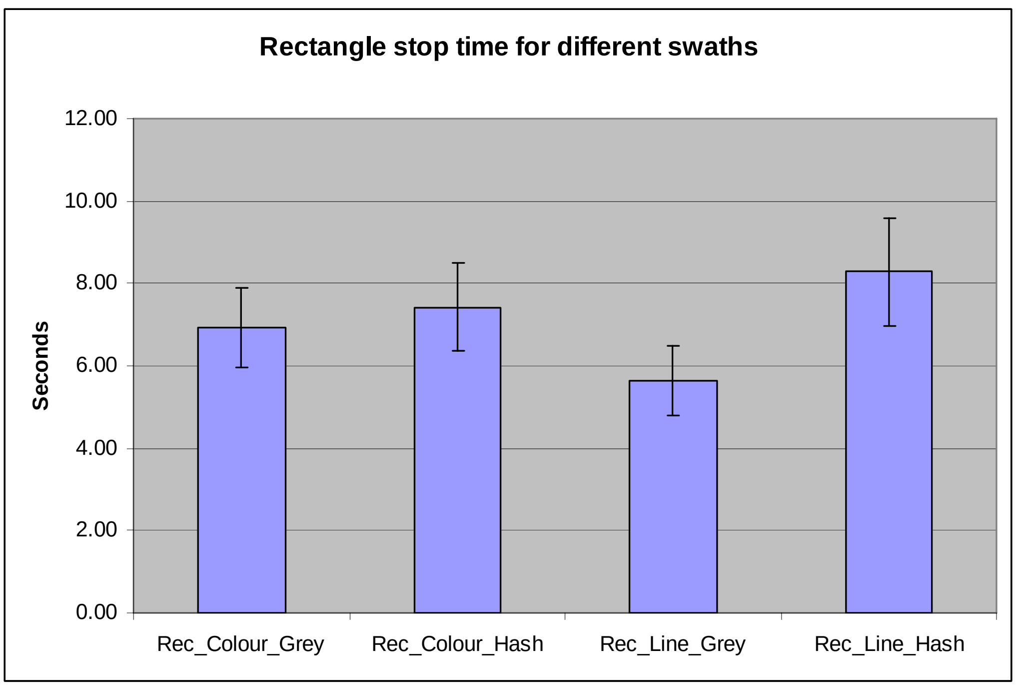



**Figure 31: Rectangle stop time for different swaths**

The following table compares the mean stop time in phases I and II. As might be expected, the stop times are shorter in phase II since the subject does not have to scan the entire screen but only the criterion swath areas and, in addition, subjects have had considerably more exposure to and practice with the target designs.

**Table 20: Comparison of "stop" time (seconds) with and without swath criteria**

| | Lego | Rectangle |
|---|---|---|
| Phase I (no swath) | 13.7 | 11.7 |
| Phase II (swaths) | 8.9 | 7.1 |

### 6.3.4 Questionnaire data

The following table summarizes the workload ratings for visualizing and finding the target using the following scale.

Use the following scale to rate the **workload** you experienced for the areas below:

1. *The load was very low and I could do this task continuously with little fatigue*
2. *The load was moderately low and I could do this task for a few hours at a time*
3. *The load was noticeable and I would need regular breaks from the task*
4. *The load was difficult but I could handle it for the duration of the experiment*
5. *The load was very high and there were times when I would have liked to stop*.

**Table 21: Workload ratings for sensor overlay conditions**

| | LEGO | | RECTANGLE | |
|---|---|---|---|---|
| Background | Grey | Hash | Grey | Hash |
| Visualizing target | 2.13 | 2.13 | 2.1 | 2.15 |
| Finding Target | 1.96 | 1.88 | 2.1 | 1.95 |

As can be seen the ratings indicated a moderately low workload on both tasks with extremely small differences between the background conditions.

Subjects also reported the presence of any untoward symptoms of visual discomfort using the following scale and the mean ratings are indicated in the following table.

1. *Not at all*
2. *Occasionally, did not bother me*
3. *Sometimes and it was quite noticeable*
4. *Frequently and it bothered me*
5. *A lot and I felt significant discomfort*

**Table 22: Visual comfort ratings for sensor overlay conditions**

| | LEGO | RECTANGLE |
|---|---|---|

| Symptom | Grey | Hash | Grey | Hash |
|---|---|---|---|---|
| Dry, itching or sore eye | 1.46 | 1.46 | 1.36 | 1.32 |
| Blinking | 1.23 | 1.29 | 1.18 | 1.23 |
| Watery eye | 1.13 | 1.13 | 1.09 | 1.18 |
| Pain around the eye area | 1.08 | 1.25 | 1.09 | 1.05 |
| Headache | 1.17 | 1.25 | 1.00 | 1.00 |
| Other muscle aches | 1.54 | 1.71 | 1.45 | 1.55 |

The very low ratings indicate that these background displays did not produce unwanted symptoms even after two hours of intensive visual search. Again, there do not appear to be any differences in ratings between the background conditions.

Given the numbers obtained, no inferential statistical tests were thought to be necessary.

# 7. Summary

## 7.1 General finding

The data show clearly that the Rectangle condition produces faster search times than the Lego condition and that both conditions produce low and equivalent error rates. The magnitude of the performance difference is approximately 16% when the search involves one or two criteria, but increases to 29% for a three criteria search. Further, whereas the Lego search times are somewhat uninfluenced by the number of search criteria, the Rectangle search time is fastest for the three criteria search. This finding suggests that subjects are able to search for multiple criteria in parallel and that possibly the overall icon becomes established as a single recognizable gestalt, instead of a combination of different elements each representing different contact properties. In phase I of the study there is evidence from the graph depicting performance for the three criteria search over trials (figure 20), that the Lego design was approaching the same speed of recognition as the Rectangle design but that the learning of the overall icon proceeded more slowly.

The “stop time” data (figure 21) also appear to support the finding that the design representations can behave as an integrated icon in that subjects were faster at concluding they had found all of the criterion targets when there were three criteria than when there was a single criterion. This could be related to the fact that subjects had more difficulty in ignoring the irrelevant search dimensions since they had learned the overall gestalt.

The “think time” data suggest that both designs are equivalent when it comes to translating the written search criteria into the to-be-searched-for representation in the “mind’s eye”.

The second phase of the experiment showed that both of the designs for sensor quality and sensor time lateness resulted in very similar search times that were comparable to the individual target search times from phase I. This suggested that the background swaths had little impact on the conspicuity of the targets. The gray area fill produced marginally (but significant) shorter target search times (.3 sec) than the fill with cross hatching. Similarly, the solid/broken line border to indicate recency of sensor coverage resulted in slightly faster search times (.38 sec) than the colored border

The subjective data showed some small advantage for the Rectangle design, however, both designs seemed to be equivalent in perceived workload and that this level was consistent with what might be expected with a visually intensive task under more operational conditions. There was no evidence from the rating data of the swath background concepts that either design was preferred. None of the designs produced unwanted symptoms of visual discomfort. Certainly, there is nothing in the subjective data which would suggest that the proposed designs have issues with workload or comfort that would prevent them being taking forward for future review by the operational community.

## 7.2 Generalizability of the results across subjects

Overall, the performance advantage for the Rectangle condition held up over all subjects, although the magnitude of the difference varied somewhat being as low as 5% for one of the fastest subjects overall and as high as 46% for one of the slowest subjects. There were also some individual differences in whether the Rectangle condition showed an advantage when the individual levels of the search criteria were separated out.

Overall, however, it would seem reasonable to conclude that the performance advantage for the Rectangle condition can be demonstrated to have generalizability across the subject population tested.

## 7.3 Magnitude of effect – operational considerations

While the performance difference between the two conditions was statistically significant, there is a question of whether a .25 to .5 second advantage for the location of a contact is of operational consequence. Certainly, if an operator were doing thousands of such searches a day, then the difference might add up to some considerable time saving. However, this is likely not the manner in which such uncertainty cues would actually be used in the operational context, where they may more likely serve a role to flag attention to certain problems in the picture. Therefore, the questions of more importance operationally may focus on how well operators believe the information is being represented, how quickly they can learn the pattern, what is the potential for errors or confusion, how do the patterns map onto existing operator population stereotypes and how disruptive might additional icons be in the general RMP.

## 7.4 Recommendations

Although the data show a consistent advantage for the Rectangle design and small advantages for the grey swath fill with a solid/broken border, it is suggested that all designs be considered candidates for presenting to the operational community.

While it is not clear that the specific mapping of time lateness onto either of the target icon designs is intuitive and without the potential for ambiguity and confusion, there was no evidence for this in the data obtained, and it is likely that training and familiarity would likely overcome such potential confusion.

Finally, it should be noted that at some future point there will need to be considerable fine tuning of the size dimensions of the prototype designs as well as the specific colors employed in order to maximize the visual and cognitive integration into the RMP display. In addition, it would be beneficial to establish with the operational community the number of levels of uncertainty that need to be represented for each of the uncertainty dimensions.

# 8. References


Cedilnik, A. and Rheingans, P. (2000), Procedural annotation of uncertain information, From the Visualization '00, IEEE Computer Society Press, Salt Lake City, Utah, United States.

Clausner, T.C. and Fox, J.R. (2005), A Framework and Toolkit for Visualizing Tacit Knowledge, In Proceedings of the 2005 International Conference on Intelligence Analysis, https://analysis.mitre.org/proceedings/index.html McLean, VA.

Davenport, M. and Franklin, S. (2006), Visions of a Future Maritime Picture: Study 1 of 3 (Limited Distribution) (DRDC Atlantic CR 2006-038;MDS-RX-RP-52-3962) Defence R&D Canada - Atlantic, Dartmouth NS (CAN);MacDonald Dettwiler and Associates Ltd, Richmond BC (CAN);MacDonald Dettwiler and Associates Ltd, Dartmouth NS (CAN).

Davenport, M. and Risley, C. (2006). Information Visualization: The State of the Art for Maritime Domain Awareness. DRDC Atlantic CR 2006-122, Defence R&D Canada Atlantic.

Finger, R. and Bisantz, A.M. (2002), Utilizing graphical formats to convey uncertainty in a decision-making task, *Theoretical Issues in Ergonomics Science*, 3 (1), 1 - 25.

Griethe, H. and Schumann, H. (2006), The Visualization of Uncertain Data: Methods and Problems, In *Proceedings of SimVis 2006*, 143–156, Magdeburg, Germany.

Pang, A.T., Wittenbrink, C.M. and Lodha, S.K. (1997), Approaches to uncertainty visualization, *The Visual Computer*, 13 (8), 370-390.

Riveiro, M. (2007), Evaluation of Uncertainty Visualization Techniques for Information Fusion, In *Proceedings of The 10th International Conference on Information Fusion (FUSION 2007),* Québec, Canada.

Stephenson, P.D., Wray, K. and Encarnação, L.M. (2006), Using Illustrative Rendering in Visualization to Intuitively Depict Information Quality and Uncertainty, In *Proceedings of Undersea Human Systems Integration Symposium*.

Unger Campbell, G. and Baker, K. (2003). Command Decision Aiding Technology (COMDAT) symbology and design study. (DRDC Toronto CR 2003-036). Unger Campbell and Associates, Langley, BC.

Ware, C. (2004) *Information Visualization: Perception for Design*. San Francisco: Morgan Kaufmann.

Xie, Z., Huang, S., Ward, M.O. and Rundensteiner, E.A. (2006), Exploratory Visualization of Multivariate Data with Variable Quality, In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology 2006*, Baltimore, Maryland. http://davis.wpi.edu/~xmdv/docs/vast06_dataqua.pdf

# 9. Acronyms

| | |
|---|---|
| AOI | Area of Interest |
| ANOVA | Analysis of Variance |
| DRDC | Defence Research and Development Canada |
| EA | Experiment Administrator |
| HREC | Human Research Ethics Committee |
| HSI® | Humansystems Incorporated |
| MDA | Maritime Domain Awareness |
| MMSI | Maritime Mobile Service Identity |
| PA | Project Authority |
| R&D | Research and Development |
| RMP | Recognized Maritime Picture |
| SOW | Statement of Work |

This page intentionally left blank.

# Annex A: List of Papers Reviewed


Pang, A. (2001), Visualizing uncertainty in geo-spatial data, From the Workshop on the Intersections between Geospatial Information and Information Technology White Papers, http://www7.nationalacademies.org/cstb/project_geospatial_papers.html, Arlington, Virginia. http://www7.nationalacademies.org/cstb/wp_geo_pang.pdf

Riveiro, M. (2007), Evaluation of Uncertainty Visualization Techniques for Information Fusion, In *Proceedings of The 10th International Conference on Information Fusion (FUSION 2007)*, Québec, Canada.

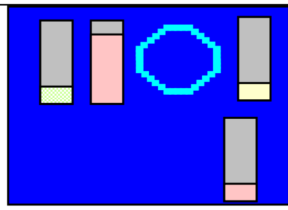
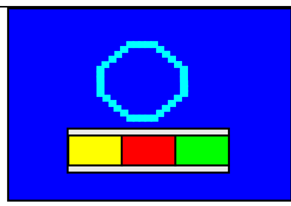
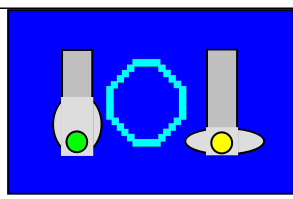
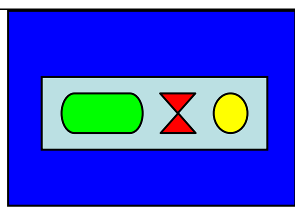
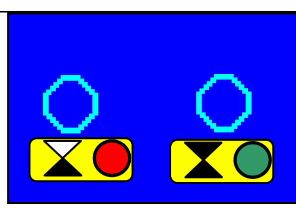
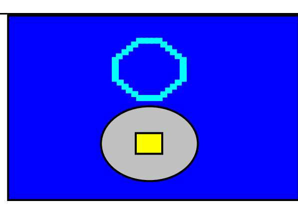
Stephenson, P.D., Wray, K. and Encarnação, L.M. (2006), Using Illustrative Rendering in Visualization to Intuitively Depict Information Quality and Uncertainty, In *Proceedings of Undersea Human Systems Integration Symposium*.

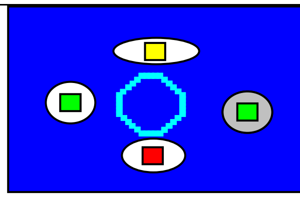
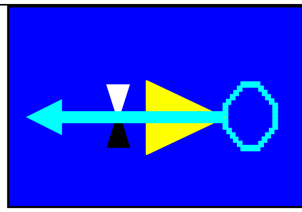
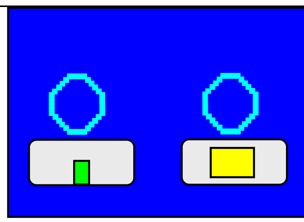
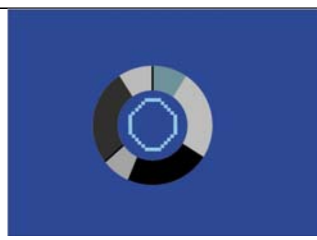
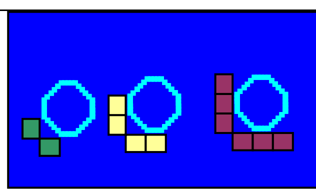
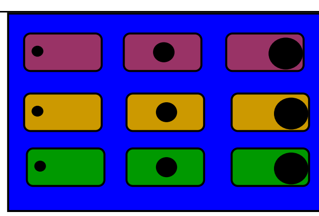
Van der Wel, F.J.M., Van der Gaag, L.C. and Gorte, B.G.H. (1998), Visual exploration of uncertainty in remote-sensing classification, *Computers & Geosciences*, 24 (4), 335-343. http://www.cs.uu.nl/research/techreps/repo/CS-1997/1997-29.pdf

Xie, Z., Huang, S., Ward, M.O. and Rundensteiner, E.A. (2006), Exploratory Visualization of Multivariate Data with Variable Quality, In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology 2006*, Baltimore, Maryland. http://davis.wpi.edu/~xmdv/docs/vast06_dataqua.pdf

This page intentionally left blank.

# Annex B: Prototype Uncertainty Designs Considered

| Number | Icon | Identity | Spatial | Time Late |
|---|---|---|---|---|
| 1 |  | Colour | X | Amount of vertical fill |
| 2 |  | Left bar color | Centre bar color | Right bar color |
| 3 |  | Circle color | Width of ellipse | Vertical height on bar |
| 4 |  | Left rectangle color | Circle color | Amount of hourglass fill |
| 5 |  | Rectangle background color | Circle color | Hourglass color |
| 6 |  | Color of square | Size of circle | Color of circle |

| Number | Icon | Identity | Spatial | Time Late |
|---|---|---|---|---|
| 7 |  | Color of square | Size of ellipse/circle | Color of circle |
| 8 |  | Color of wedge | Angular width of wedge | Hourglass fill |
| 9 |  | Color of small rectangle | Width of small rectangle | Height of small rectangle |
| 10 |  | Amount/color of sector fill | Amount/color of sector fill | Amount/color of sector fill |
| 11 |  | Color of square segment | Number of horizontal segments | Number of vertical segments |
| 12 |  | Colour of background | Size of circle | Position of circle |

# Annex C: Approved Ethics Protocol

## Executive summary

**Title:** Comparing Techniques for Visualizing Uncertainty

**Principal Investigator:** Dr. Michael Matthews, Humansystems Incorporated® (HSI®). Tel: 519-836-5911

**Defence Reasearch and Development Canada (DRDC) Co-Investigators:**

Ms. Liesa Lapinski, DRDC Atlantic. Tel: 902-426-3100 x180

Ms. Sharon McFadden, DRDC Toronto. Tel: 416-635-2189

**Thrust:** 11he Maritime Domain Awareness

**Objectives:**

The goal of this experiment is to explore, in the context of the Recognized Maritime Picture (RMP), ways of visualizing the temporal, spatial, and identity uncertainty of vessel tracks as well as the coverage uncertainty of the sensors and information sources. Temporal uncertainty is defined in terms of how stale the information is on a vessel; spatial uncertainty is not knowing the location precisely and identity uncertainty concerns not knowing information such as the vessel's name or hull number.

**Overview:**

Twenty five volunteer subjects will be required to conduct two computer assisted tasks using a simulation of the RMP. The tasks will involve the volunteers scanning a screen that contains a number of symbols and locating those that meet criteria established at the beginning of each trial run. A questionnaire will be administered at the end of each session, to record the volunteers' subjective evaluation of the different display techniques (e.g. use of colour and symbols) used. The experiment will be undertaken in Guelph, Ontario at Humansystems, Inc. and will include 8 test sessions. Each session will comprise 25 minutes of training, a 10 minute break and 50 minutes of testing. Participants may choose to schedule multiple sessions (up to maximum of 3) per day with a minimum of a 30 min break between sessions, or to spread sessions across days, at their own convenience. This work will be conducted by Humansystems Incorporated.

**Subjects**

Male or female volunteer subjects between the ages of 18-55 will be recruited from the local community and paid for their participation.

**Risks**

This experiment offers minimal risk to the participant's health and well-being. There is a low risk of eye fatigue or eyestrain, as is associated with doing any visually intensive task on a computer display. Any potential risks will be mitigated by giving participants regular breaks in the task.

## Protocol # L-631

**Title:** Comparing Techniques for Visualizing Uncertainty

**Principal Investigator:**

Dr. Michael Matthews, Humansystems Incorporated® (HSI®). Tel: 519-836-5911

**DRDC Co-Investigators**

Ms. Liesa Lapinski, DRDC Atlantic. Tel: 902-426-3100 x180

Ms. Sharon McFadden, DRDC Toronto. Tel: 416-635-2189

**Thrust:** 11he Maritime Domain Awareness

**List of Acronyms**

| | |
|---|---|
| DRDC | Defence Research and Development Canada |
| HSI® | Humansystems Incorporated |
| MDA | Maritime Domain Awareness |
| RMP | Recognized Maritime Picture |
| RJOCs | Regional Joint Operations Centres |

**Background:**

This applied research project in the Maritime Domain Awareness (MDA) Thrust is studying information visualization and management for enhanced domain awareness in maritime security. The DRDC/HSI® team want to investigate the best way to visualize particular aspects of uncertainty and then test to see if the visualizing methods can help improve understanding of the Recognized Maritime Picture (RMP), decision making based on the RMP and the efficiency of the RMP operators' duties.

The RMP is a product produced by the Regional Joint Operation Centres (RJOCs). In its common form, it is a map of the Canadian coastal waters, with contacts, typically ships, marked on the map. Each contact has a set of metadata associated with it which can include (but is not exclusive to) position, speed, heading, ship name, hull number, threat, flag, destination, origin, type, cargo and a digital image. At worst, the metadata only consist of a position (i.e., there is something out there). At best, the metadata consist of all of the above. The different degrees of metadata are due to the multiple sources of information that feed the RMP. These sources include everything from radar to surveillance flights to self reporting systems to voluntary reports, each providing its own subset of data.

Uncertainties the RMP operators and users may face in creating the RMP can arise from, for example:

- the metadata provided by a source being grossly out of date, and/or wrong, and/or missing;
- a sensor often providing false contact reports, with no indication of when it is false or not;
- the time-latency associated with each contact; and,
- contacts not being present, which can make it unclear whether there is an absence of a contact or the area hasn't been surveyed recently

The purpose of the proposed experiment is to explore the best ways of representing these uncertainties. Specifically, we are interested in visualizing the temporal, spatial, and identity uncertainty of each track and the coverage uncertainty of the sensors and information sources of the RMP. Temporal uncertainty is defined in terms of how stale the information is on a vessel; spatial uncertainty is not knowing the location precisely and identity uncertainty concerns not knowing information such as the vessel's name or hull number. The outcome of this work will serve both the maritime operations communities, as well the scientific communities in the advancement of new uncertainty visualizations. This work will be conducted by Humansystems Incorporated.

**Objectives:**

The goal of this experiment is to explore, in the context of the Recognized Maritime Picture (RMP), ways of visualizing the temporal, spatial, and identity uncertainty of tracks as well as the coverage uncertainty of the sensors and information sources. This work will be conducted by Humansystems Incorporated.

**Overview:**

Twenty five male or female volunteer subjects will be required to conduct a computer assisted task using a simulation of the RMP. The task will involve the volunteers looking through a number of symbols that have different levels of uncertainty represented by symbols or colour coding and selecting those that meet criteria established at the beginning of each trial run. A questionnaire will be administered at the end of each session, to record the volunteers' subjective evaluation of the different methods used to represent uncertainty. The experiment will be undertaken in Guelph, Ontario at Humansystems, Inc. and will include 8 test sessions. Each session will comprise 25 minutes of training, a 10 minute break and 50 minutes of testing. Participants may choose to schedule multiple sessions (up to maximum of 3) per day with a minimum of a 30 min break between sessions, or to spread sessions across days, at their own convenience.

**Procedures:**

The basic task

The subject is presented with a visual display that is designed to simulate a simplified version of the RMP. This screen includes a number symbols, ranging between 10-50, representing different contacts (vessels) with different uncertainties associated. These symbols will include uncertainty information that is represented in different formats (e.g. variations in colour hue, saturation and different shapes or symbols (e.g. ellipses, icons[8]). Specific values of the colour or symbols will be associated with different levels of uncertainty for different vessel tracks and sensor characteristics. These visual characteristics may be supplemented with text and numerical annotations.

In addition, to the representation of uncertainty in the contacts, area fill techniques for representing sensor coverage will also be explored. Methods such as colour (saturation and intensity) and area fill (e.g. hashing) will be compared.

[8] See Annex D for samples of coding that are being evaluated for implementation in the experiment.

Subjects will be required to locate contacts by searching the screen to identify vessel tracks that meet levels of uncertainty prescribed at the start of each trial run. For example, in one trial run subjects would “de-clutter” the picture, where the objective would be to find and delete some specified type of contacts (e.g. those contacts where sensor coverage is beyond some “uncertainty” threshold). For this task, the subjects start with the simplified RMP screen showing 10-50 contacts with varying levels of uncertainty. The subjects would then use the mouse to click on targets that were considered “uncertain” and these contacts would disappear. The subjects would continue until they believed the RMP was “de-cluttered”. To show they were complete, the subjects would use the mouse to click on a “Finished” button in the lower section of the screen.

A second task will require subjects to select contacts that meet a particular uncertainty criterion, for example, “find all contacts which have been reported in the last 20 minutes, for whom the position is known within in a 10 nautical mile radius, where radar coverage is strong”.

Both types of task will be performed by all subjects in separate trial blocks with the order counterbalanced across subjects.

Two coding schemes will be evaluated for each the three types of contact uncertainty and to depict the sensor coverage uncertainty to yield eight experimental treatments. Subjects will be trained in each coding scheme prior to a test run with that scheme.

The timing of the subject’s responses will be recorded as will the accuracy of their selections in meeting the prescribed uncertainty criteria.

At the completion of each of the testing sessions the subjects will complete a subjective questionnaire documenting their perceptions of the different visualizations of uncertainty. Annex A contains example questions of the questionnaire.

Orientation Briefing: Prior to training, participants will be given an orientation briefing on the overall study, its objectives and test activities. At this stage they will be asked to complete a consent form. (Annex C)

Training Session: Following the orientation and prior to the start of any testing, participants will be provided with a 25 minute training session. Simplified displays with variations of the uncertainty encoding method will be used to ensure that subjects have learned how the specific colour and shape coding have been mapped onto the varying levels of uncertainty. Feedback would be provided in the learning trials to ensure subjects reach the appropriate criterion level of performance.

Subjects will be required to perform the task of identifying varying levels of uncertainty to a criterion level of performance (no matter how long they take) before they are allowed to proceed with the testing session. Subjects who are unable to do the task without error will be thanked for their participation, paid and dismissed from further involvement in the study.

Testing Sessions: Following the training session, participants will have a 10 min break and then be required to perform a 50 minute test session. In all, there will be 8 test sessions arranged at the convenience of the participants. Subjects may opt to do more than one session per day (up to a maximum of 3) with a required break of 30 minutes between sessions. The procedures for each investigation will be conducted in a standardized, repeatable protocol under controlled laboratory conditions.

The following procedures will be common to all of the lab tests.

Orientation and training are provided for the particular coding method for the test session (25 minutes)

Ten minute break

Participants will be reminded of the experimental protocol prior to beginning the test.

From the "GO" instruction the participant will perform the test for 50 minutes. Within this session, there will be a number of trial blocks each dedicated to one of the experimental tasks outlined above. If the subject is participating in a second test session they will then have a 30 minute break.

At the completion of each test sessions the participants will complete a questionnaire about their experiences with the different visualizations (5 minutes).

Following the completion of the questionnaire on the final testing session, the subjects will have the opportunity to ask questions and provide any additional feedback.

**Participants:**

Approximately, twenty five male or female volunteers will be recruited from the local university or community to participate in this experiment. They will be recruited through an unpaid advertisement in the student newspaper and/or a poster placed upon University Bulletin Boards (See Appendix B) and possibly through Craig's list. Twenty subjects are required, so the preliminary selection of twenty five allows for some voluntary attrition or those who fail to meet training criteria, outlined above. Participants should be in the 18-55 age range to be representative of the target population of operators who work with RMP displays. Participants will be required to self identify that they have normal colour vision.

**Equipment and Facilities:**

The apparatus comprises a standard "Windows" workstation with 19" colour screen, a mouse and keyboard input, a work surface to record notes, and an ergonomically designed operator's chair.

**Data collected**

The following information will be collected during each testing session:

- The task requirements (e.g. de-clutter) as well as the associated uncertainty criteria
- Which contacts are selected
- The time of each selection
- Total time for each task
- Responses to the questionnaires concerning the subjects subjective experiences with the different tasks and types of uncertainty codings

**Experimental Design/Statistical Analysis:**

In order to maximize sensitivity of the experiment in detecting expected effects, a within-subjects design will be used, in which all subjects participate in all 8 coding methods for uncertainty. A multi-factor, repeated measures analysis of variance design will be used to evaluate the response time and uncertainty data to provide an estimate of the effectiveness of the different uncertainty visualizations. Response time data may be normalized if the underlying distribution is found to be skewed. Specific planned comparisons among experimental conditions will be evaluated. The threshold for establishing the acceptable error for statistical testing will be determined in discussions with the Scientific Authority.

## Risks and Safety Recommendations:

This experiment offers minimal risk to the participant's health and well-being. There is a low risk of eye fatigue or eyestrain, as would be associated with doing any visually intensive task (e.g. web searching, word processing) on computer display for the period of time used in the test sessions. This may manifest itself as eye discomfort, dry or itchy eyes, or mild headache. However, the duration of exposure to the visual displays will be short and the conduct of testing across separate days, or providing 30 minute breaks between sessions on a single day, will mitigate this risk. Participants will be encouraged to inform experimenters if they experience any discomfort or eyestrain, or if they have any problems during the investigation. They may be told to stop their activities until problems or conditions are resolved. There is no requirement for medical screening or medical coverage for this study.

The risks from participation in this experiment are generally the same as those associated with the performance of normal monitoring of a visual display that a person might do while word processing, surfing the web or playing video games.

**Benefits of Study:**

Through their involvement in the study subjects will be able to contribute to the validation, development and design of new methods for the coding of uncertainty, which in turn provides important human factors data for the re-design and automation of future systems to represent the maritime picture.

**Informed Consent:**

Subjects will be fully briefed on the relevant aspects of the experimental protocol and will be given a copy of this protocol to review. They will be required to sign a voluntary consent form (attached as Appendix C to this Protocol), indicating their willing informed consent, before being allowed to participate in the experiment.

**Confidentiality:**

Any personal or performance data collected for each subject will be available only to the experimenters and will be held in the strictest confidence. Subjects will not be identified by name in the data records; group statistics will be used in future presentations or publications. Individual subject data will be coded anonymously and maintained in a computer file. The file may be accessed only by the project team.

**Subject Debriefing:**

Subjects will be permitted to ask any questions they wish about the study after they have completed the experiment.

**Subject Remuneration:**

Subjects will be paid subject pay in the amount of between $149.65 and $175.44 for participation in this study, depending upon the number of sessions they choose to do in a day (in accordance with DRDC guidelines and as authorized by DND policies[9]). Partial remuneration may also occur if for any reason all sessions are not completed. The remuneration is calculated as follows:

A. Eight separate daily sessions:

2 hours x 2 (stress level) x $2.50 +$11.93 x 8 days = $175.44

B. Two days with 3 sessions per day, plus one day with two sessions per day.

6 hours x 2 (stress level) x $2.50 +$11.93 x 3 days **PLUS**

4 hours x 2 (stress level) x $2.50 +$11.93 x 2 days = $149.65

**Approximate Time Involvement:**

*All subjects will be required to participate in 8 x 85 minute sessions. These may be distributed across days or up may be scheduled (up to 3 sessions) on a single day. The total time commitment for participation in this study will therefore be approximately 11.3 hours.*

**Attachments:**

Appendix A: Sample Questionnaire

Appendix B: Recruitment advertisement and poster

Appendix C: Informed Voluntary Consent Form

[9] Guide to Stress Compensation for Human Subjects, R. Pigeau. DRDC Toronto.

## Appendix A

### Sample Questionnaire

**Research Project Title:** Comparing Techniques for Visualizing Uncertainty

**Instructions:** Please respond to each statement below by selecting the most appropriate rating response based on your experiences today.

**Ratings:**
1 – Strongly Disagree
2 – Disagree
3 – Neutral
4 – Agree
5 – Strongly Agree

**Example Questions:**

| # | Statement | Response | | | | |
|---|---|---|---|---|---|---|
| 1 | The level of uncertainty was easy to comprehend. | 1 | 2 | 3 | 4 | 5 |
| 2 | The graphic symbols were easy to distinguish from the background. | 1 | 2 | 3 | 4 | 5 |
| 3 | The symbols representing the individual levels of uncertainty were clearly differentiated. | 1 | 2 | 3 | 4 | 5 |
| 4 | I could easily determine the level of uncertainty from the symbol. | 1 | 2 | 3 | 4 | 5 |
| 5 | The symbol representing the composite level of uncertainty was easy to comprehend. | 1 | 2 | 3 | 4 | 5 |
| 6 | It was easy to tell when the information was out of date. | 1 | 2 | 3 | 4 | 5 |
| 7 | The colours used to represent the different sensor sources were easy to distinguish. | 1 | 2 | 3 | 4 | 5 |
| 8 | I could easily determine the sensor source and its level of uncertainty from the symbol. | 1 | 2 | 3 | 4 | 5 |
| 9 | The colour intensity and saturation representing the sensor coverage was easily understood. | 1 | 2 | 3 | 4 | 5 |
| 10 | The uncertainty coding scheme was easy to remember. | 1 | 2 | 3 | 4 | 5 |

## Appendix B: Recruitment advertisement and poster

### Advertisement:

Male or female research participants with normal colour vision are required to perform a task that involves visually scanning and screen containing symbols and finding the location of certain symbols using a computer display. There will be 8 sessions of approximately 85 minutes which can be scheduled at your convenience. Up to 3 sessions may be performed on a day. Volunteers should have no vision problems that would prevent them from closely watching a computer screen for up to 75 minutes at a time. If you are between the ages of 18-55, please contact Jeff Bennett at Human*systems* Incorporated at 519-836-5911 who will provide you with full details of what will be required.

**CALL-FOR-SUBJECTS POSTER**

**Title of Experiment:**

Comparing Techniques for Visualizing Uncertainty

**Purpose of Experiment**:

The experiment is designed to explore the best ways of visualizing uncertainty

**Types of Subject Requested**:

Twenty-five male or female subjects between the ages of 18-55. Subjects should not have any vision problems that would cause them to have undue discomfort when looking at a computer display for two hours. Subjects should have normal colour vision.

**Procedures**:

Subjects will be trained in a visual identification task that requires locating contacts and identifying different levels of uncertainty. The task involves viewing a number of colored symbols coded with different levels of uncertainty against a black background screen. The task requires some patience and accuracy and will involve interacting with a computer display using a mouse for a 75 minute viewing period.

**Location of Experiments:**

Testing will take place at Humansystems Incorporated offices, 111 Farquhar St. Guelph, during business hours at times that are convenient to the participant.

**Invasive Procedures and Non-Invasive Measurements Required**:

No invasive procedures are required. Subjects will be asked to fill in a questionnaire with ratings scales. Decisions made as well as the timing of those decisions will be recorded.

**Duration of Subject Participation:**

There will be eight sessions comprising 25 minute training a 10 minute break and a 50 minute test session. Participants may schedule the sessions on different days or up to 3 sessions per day.

**Risks to Subject:**

This experiment offers minimal risk to the participant's health and well-being. There is a low risk of eye fatigue or eyestrain, as is associated with any computer display use over two hours. This may manifest itself as eye discomfort, dry or itchy eyes, or mild headache

The risks from participation in this experiment are generally the same as those associated with the performance of any visually intensive task (e.g. surfing the web, word processing) on a computer display.

**Benefits:**

Individual subjects will gain insight into how Human Factors Research is conducted. There is a more general benefit to the scientific community in finding ways to visualize uncertainty.

**Compensation:**

Participants will be compensated according to DRDC guidelines and will receive between $149.65 and $175.44 depending upon how they distribute their test sessions across days.

**Point of Contact:**

Please contact Jeff Bennett at Humansystems Inc. Telephone 519-836-5911 during normal business hours.

**Appendix C**

**Voluntary Consent Form**

**Protocol Number: L-631**

**Research Project Title: Comparing Techniques for Visualizing Uncertainty**

**Principal Investigator:**

Dr. Michael Matthews,

**DRDC Co-Investigators:**

Ms. Liesa Lapinski.

Ms. Sharon McFadden

*1.* I,_________________________________________________________
_______________________________________________________________

(Name, Address, Phone number) hereby volunteer to participate as a test subject in the experiment entitled "Comparing Techniques for Visualizing Uncertainty", the aim of which is to explore the best ways of visualizing uncertainty of vessel tracks in a maritime context. I understand that I am required to read the attached protocol in its entirety. I have had the opportunity to study and discuss the attached protocol with the investigators and I have been informed to my satisfaction about the possible discomforts associated with these tests.

2. I am aware that I will participate in a 25 minute training session to learn how to use the software, and after a 10 break, a testing session of 50 minutes. Thus, the total involvement will be 85 minutes per session for 8 test sessions. I may choose to do up to 3 sessions per day. During each session I will receive a briefing on the aims and procedures for the experiment. I will have the opportunity to ask and receive answers to any questions I may have. After each test session, I will be given at least a 30 minute rest period before starting the next session. I understand that I am free to refuse to participate and may withdraw my consent without prejudice or hard feelings at any time. Should I withdraw my consent, my participation as a subject will cease immediately.

3. I have been told that the principal risks associated with this experiment are the possible development of eyestrain or visual fatigue. This may manifest itself as eye discomfort, dry or itchy eyes, or mild headache. I understand that the limited duration of each experiment and rest periods between experiments will mitigate this risk. I understand and accept this risk. I am aware that there are inherent, unknown and currently unforeseen risks by DRDC Atlantic and the Project Investigators that are associated with any scientific research and that all known risks have been explained to my satisfaction.

4. I have been given examples of potential minor and remote risks associated with the experiment and consider these risks acceptable as well. Also I acknowledge that my participation in this study, or indeed any research, may involve risks that are currently unforeseen by DRDC Atlantic.

5. I agree to provide responses to questions that are to the best of my knowledge truthful and complete. I have been advised that the experimental data concerning me will be treated as confidential and not revealed to anyone other than the investigators without my consent except as

data unidentified as to source. I am aware that there will not be any on-site medical coverage during the experiment. I understand that my name will not be identified or attached in any manner to any publication arising from this study.

6. I understand that for my participation in this research project, I am entitled to remuneration in the form of subject payment of between $149.65 and $175.44, depending on the number of sessions I choose to do in one day. Stress remuneration is taxable. However, T4A slips are issued only for amounts in excess of $500.00 remuneration per year.

7. I acknowledge that I have read this form and I understand that my consent is voluntary and has been given under circumstances in which I can exercise free power of choice. I have been informed that I may, at any time, revoke my consent and withdraw from the experiment, and that the investigators may terminate my involvement in the experiment, regardless of my wishes.

8. I understand that by signing this consent form I have not waived any legal rights I may have as a result of any harm to me occasioned by my participation in this research project beyond all risks I have assumed.

Volunteer's Name: ______________________________________________

Signature: ___________________________ Date: ____________________

Name of Witness to Signature: ___________________________________

Signature: ___________________________ Date: ____________________

Family Member or Contact Person (name, address, daytime phone number & relationship):

______________________________________________________________

______________________________________________________________

Principal Investigator: ___________________________________________

Signature: ___________________________ Date: ____________________

FOR SUBJECT ENQUIRY IF REQUIRED:

Should I have any questions or concerns regarding this project before, during, or after participation, I understand that I am encouraged to contact any of the contacts below by phone or e-mail, to

**Principal Investigators:**

Dr. Michael Matthews, Tel: 519-836-5911, email: mmatthews@humansys.com

Co-Investigator:

Liesa Lapinski, Tel: 902-426-3100 x180, email: Liesa.Lapinski@drdc-rddc.gc.ca

Chair, DRDC Human Research Ethics Committee (HREC): Dr. Jack P. Landolt, phone: 416-635-2120, email: jack.landolt@drdc-rddc.gc.ca

I understand that I will be given a copy of this consent form so that I may contact any of the above-mentioned individuals at some time in the future should that be required.

Secondary Use of Data: I consent/do not consent (delete as appropriate) to the use of this study's experimental data involving me in unidentified form in future related studies provided review and approval have been given by DRDC HREC.

Volunteer's Signature_____________________ Date ____________________

. This page intentionally left blank.

# Annex D: Workload and Visual Comfort Rating Scales

Rate your **agreement** with the following statements, using the following scale:

1. Completely disagree
2. Somewhat disagree
3. Neither agree or disagree
4. Somewhat agree
5. Completely agree

| # | Statement | Response | | | | |
|---|---|---|---|---|---|---|
| 1 | The different levels of identity information were easy to comprehend. | 1 | 2 | 3 | 4 | 5 |
| 2 | The different levels of the spatial information were easy to comprehend. | 1 | 2 | 3 | 4 | 5 |
| 3 | The different levels of the time information were easy to comprehend. | 1 | 2 | 3 | 4 | 5 |
| 4 | The visual representations of the different levels of the identity information were easy to discriminate. | 1 | 2 | 3 | 4 | 5 |
| 5 | The visual representations of the different levels of the spatial position information were easy to discriminate. | 1 | 2 | 3 | 4 | 5 |
| 6 | The visual representations of the different levels of the time information were easy to discriminate. | 1 | 2 | 3 | 4 | 5 |
| 7 | It was easy to know what to look for when there were multiple criteria | 1 | 2 | 3 | 4 | 5 |
| 8 | It was easy to find the contacts when there were multiple criteria | 1 | 2 | 3 | 4 | 5 |

Use the following scale to rate the **workload** you experienced for the areas below:

1. The load was very low and I could do this task continuously with little fatigue
2. The load was moderately low and I could do this task for a few hours at a time
3. The load was noticeable and I would need regular breaks from the task
4. The load was difficult but I could handle it for the duration of the experiment
5. The load was very high and there were times when I would have liked to stop.

| # | Statement | Response | | | | |
|---|---|---|---|---|---|---|
| 1 | The task of thinking about what to look for based on the search criteria | 1 | 2 | 3 | 4 | 5 |
| 2 | The task of visually trying to find each contact | 1 | 2 | 3 | 4 | 5 |

Using the following scale, rate your experience with any of the following symptoms of **visual discomfort** during the course of the experiment

1. Not at all
2. Occasionally, did not bother me
3. Sometimes and it was quite noticeable
4. Frequently and it bothered me
5. A lot and I felt significant discomfort

| # | Statement | Response | | | | |
|---|---|---|---|---|---|---|
| 1 | Dry, itching or sore eye | 1 | 2 | 3 | 4 | 5 |
| 2 | Blinking | 1 | 2 | 3 | 4 | 5 |
| 3 | Watery eye | 1 | 2 | 3 | 4 | 5 |
| 4 | Pain around the eye area | 1 | 2 | 3 | 4 | 5 |
| 5 | Headache | 1 | 2 | 3 | 4 | 5 |
| 6 | Other muscle aches | 1 | 2 | 3 | 4 | 5 |

# Annex E: Sample Training Materials

The following sample is based upon a PowerPoint interactive presentation given to experiment participants at the start of each session. The particular example provided was from Phase II of the experiment involving the search for contacts against swath areas depicting sensor coverage. In this Phase, participants did not receive detailed training on the search for the actual contacts, as this had been trained in Phase I. Instead, they were provided with a "refresher" of the principles involved.

# Welcome to Human*systems* Inc

Welcome to Human*systems* Inc. Thank you for agreeing to participate in this study.

This PowerPoint presentation will introduce you to a set of visualizations which represent information about the quality and timeliness of coverage of an area of sea on two separate dimensions. It will also review the symbols you are already familiar with that relay information on contacts.

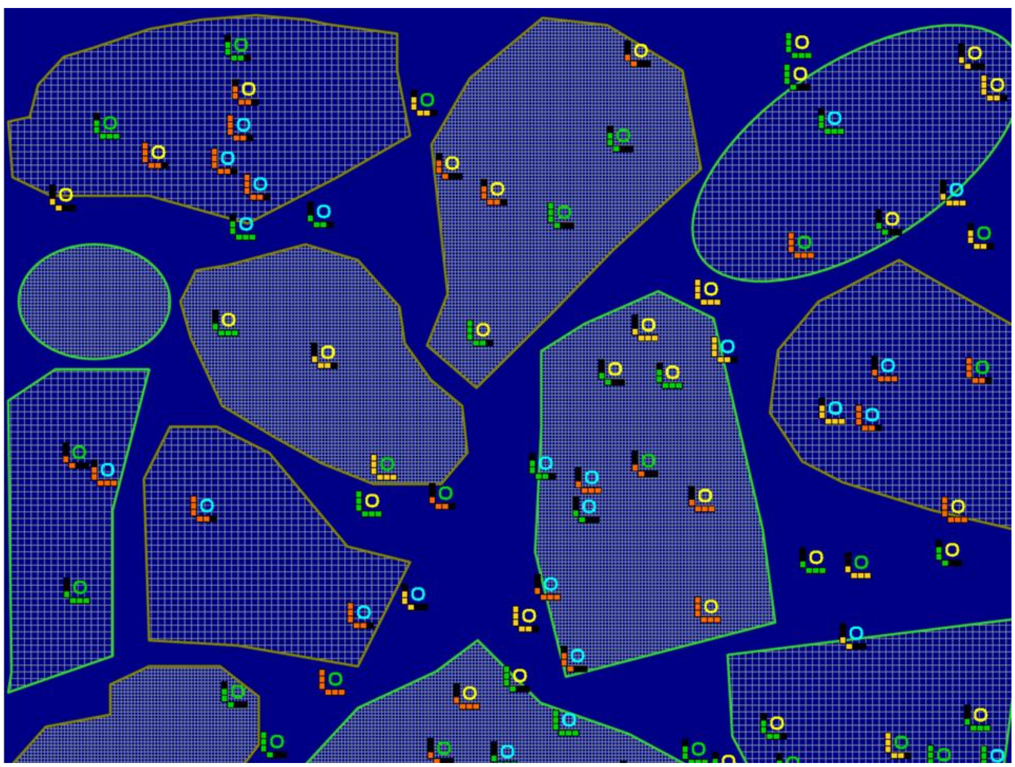
The objective of this session is to familiarize you with what the coverage looks like so that you will be able to readily recognize them, interpret the information that they convey, and be able to select the appropriate symbols. The following slide shows you how the symbols will appear in the context of the experimental study. The next slide is for familiarization purposes only. Please take a look and then move on to the next slide when you are ready.

At the end of this training session there is a short test.



Integrating People, Performance & Technology™

2



# Introduction & outline

- The type of border and fill of the different coloured regions that you see represent the sensor coverage of that area of the ocean
- Each coverage region is coded with the following depending on the fill and border of the coloured region;
  - **Quality Information** (how reliable the data are from that sensor)
  - **Time Information** (how recent data from the area is)



Integrating People, Performance & Technology™

4

# Symbol design

Before we get into the new information about these coloured regions…

- Lets quickly review the symbols representing information about different contacts (which you are already familiar with)

- Each contact symbol is rated using the following terminology

  IDENTITY INFORMATION: ***Good/Poor/None***

  TIME INFORMATION: ***Current/Old/Out of date***

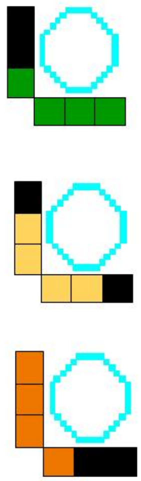
  SPATIAL POSITION: ***Precise/Approximate/Unknown***

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

5

# General introduction



- The quality of **IDENTITY** information relating to a vessel is indicated by *colour*
  - **Green** = **Good** Identity Information
  - **Yellow** = **Poor** Identity Information
  - **Orange** = **No** Identity Information
- The quality of **TIME** information is indicated by the number of blocks in the *vertical bar*
  - **One Block** = **Current** Time Information
  - **Two Blocks**= **Old** Time Information
  - **Three Blocks** = **Out of Date** Time Information
- The quality of **SPATIAL** information is indicated by number of blocks in the *horizontal bar*
  - **One Block** = **Precise** Spatial Location
  - **Two Blocks** = **Approximate** Spatial Location
  - **Three Blocks** = **Unknown** Spatial Location
- Now, we will get back to the coverage regions and the types of information associated with those.



Integrating People, Performance & Technology™

6

# QUALITY information

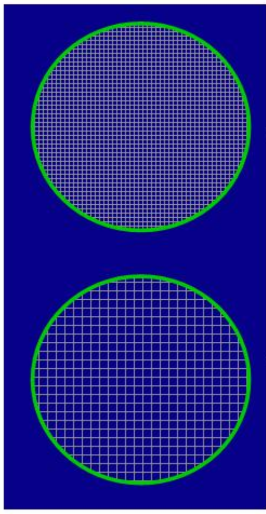
- **QUALITY INFORMATION** refers to how accurate the indicated coverage is
  - Can be either "***strong***" ***or*** "***weak***"
- **QUALITY INFORMATION** is represented by the ***fill*** of the coverage area
- Please move on to the next slide; this concept is explained in more detail there;



Integrating People, Performance & Technology™

7

# QUALITY



The QUALITY of the coverage of the region is indicated by the coverage region's **fill**

- If the QUALITY of the coverage is ***strong***, then the fill is ***fine grid***
- If we have ***weak*** QUALITY coverage information, then the fill will be ***coarse grid***



Integrating People, Performance & Technology™

8

# Questions

- How many levels of QUALITY coverage are there?
- Which fill indicates that ***strong*** QUALITY coverage is available?
- If the fill is a coarse grid what does this mean?

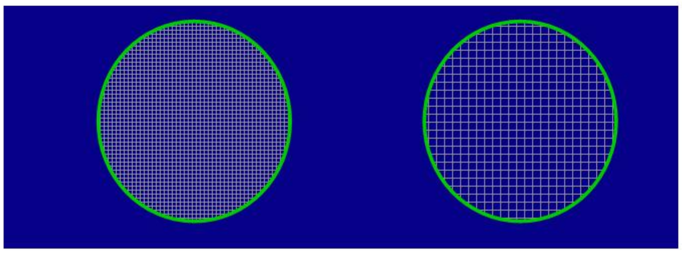


Integrating People, Performance & Technology™

9

# Questions

- Which symbol below means that there was ***weak*** QUALITY coverage?



©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

10

# TIME information

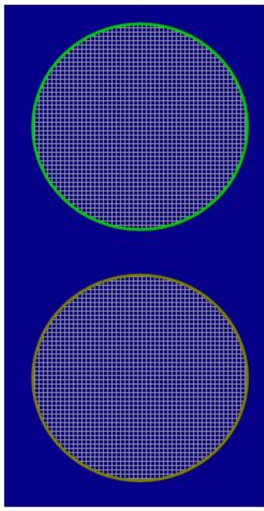
- **TIMING** of the coverage region refers to how up to date the indicated coverage is
  - Can be either "***current" or "out of date***"
- **TIMING INFORMATION** is represented by the border of the coverage area
- Please move on to the next slide; this concept is explained in more detail there;

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

11

# TIME



The **TIMING** information about the coverage region is shown through the **border**

- If TIME information is ***current***, then the border is ***green coloured***
- If we have ***out of date*** TIME information, then the border will be ***gold coloured***

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

12

# Questions

- How many levels of TIMING information are there for coverage regions?
- Which border indicates that the coverage is ***current***?
- If the border is ***gold coloured*** what does this mean?

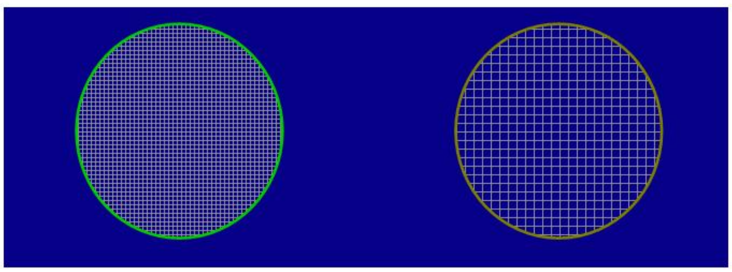
©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

13

# Questions

- Which symbol below means that there was ***out of date*** TIMING information?



©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

14

# Questions

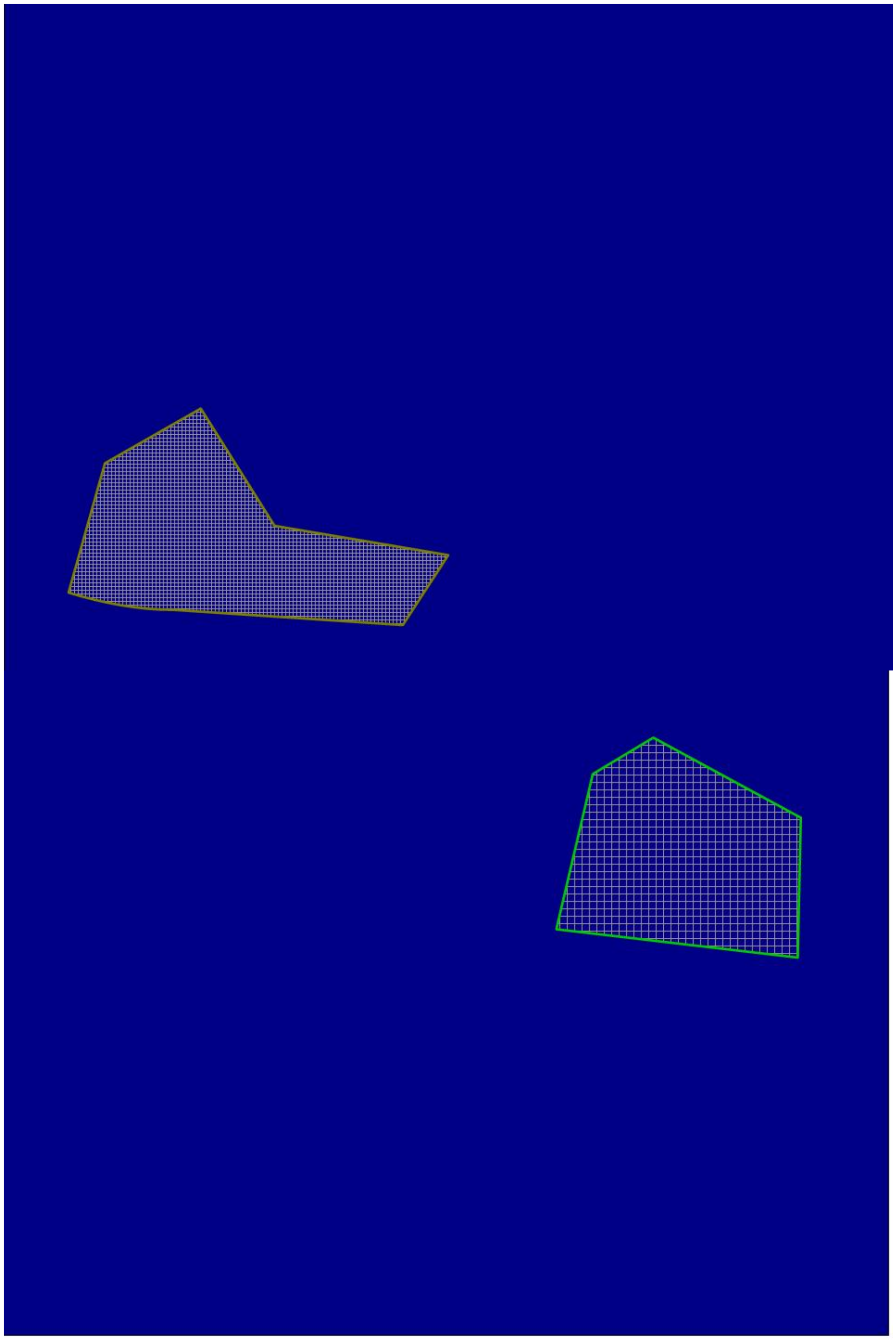
The following slides show different coverage areas. Please describe the level of:

- QUALITY – ‘Strong’ or ‘Weak’
- TIME – ‘Current’ or ‘Out of Date’

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

15



## How would you represent the following coverage?

- ***Weak*** QUALITY coverage information
- ***Out of date*** TIME coverage information

©Human*systems*® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

18

## How would you represent the following coverage?

- ***Strong*** QUALITY coverage information
- ***Current*** TIME coverage information

©Human*systems*® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

19

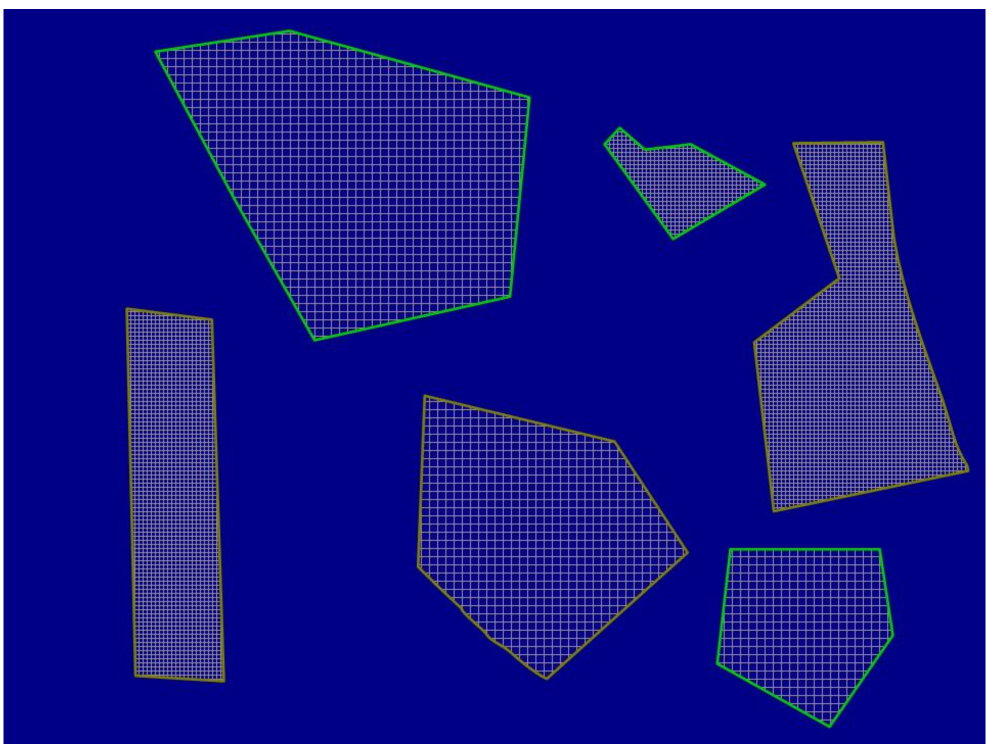
# Task 1

- The following slide displays an array of coverage types
- Please use your mouse to move the cursor over and point out the coverage regions that are CURRENT

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

20



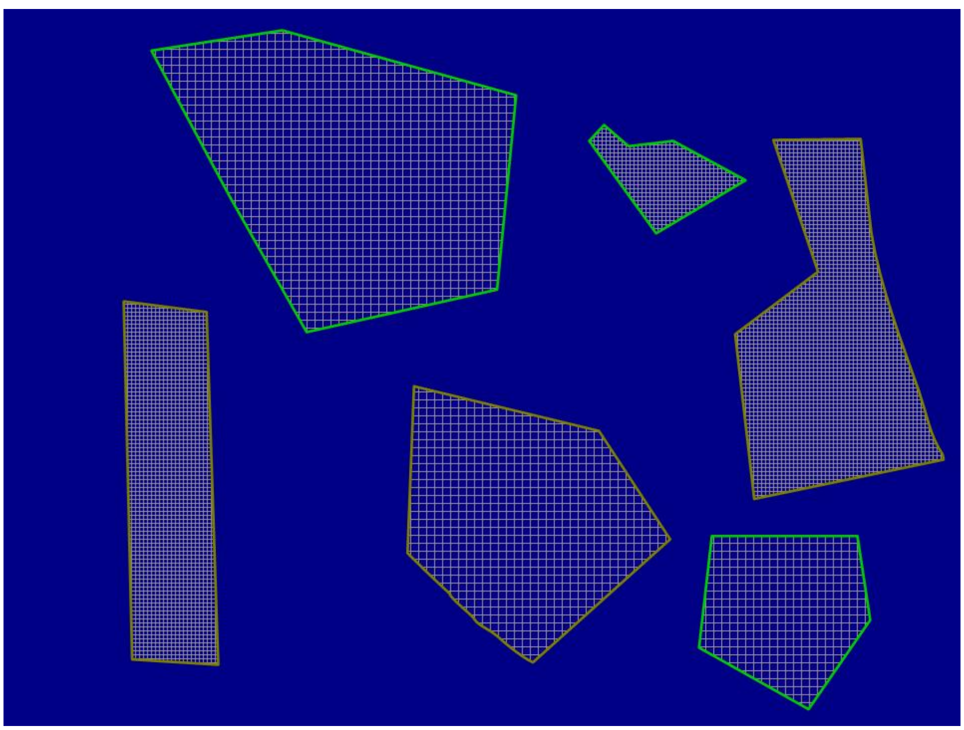
# Task 2

- Please use your mouse to move the cursor over and point out the coverage regions that have STRONG coverage quality



Integrating People, Performance & Technology™

22



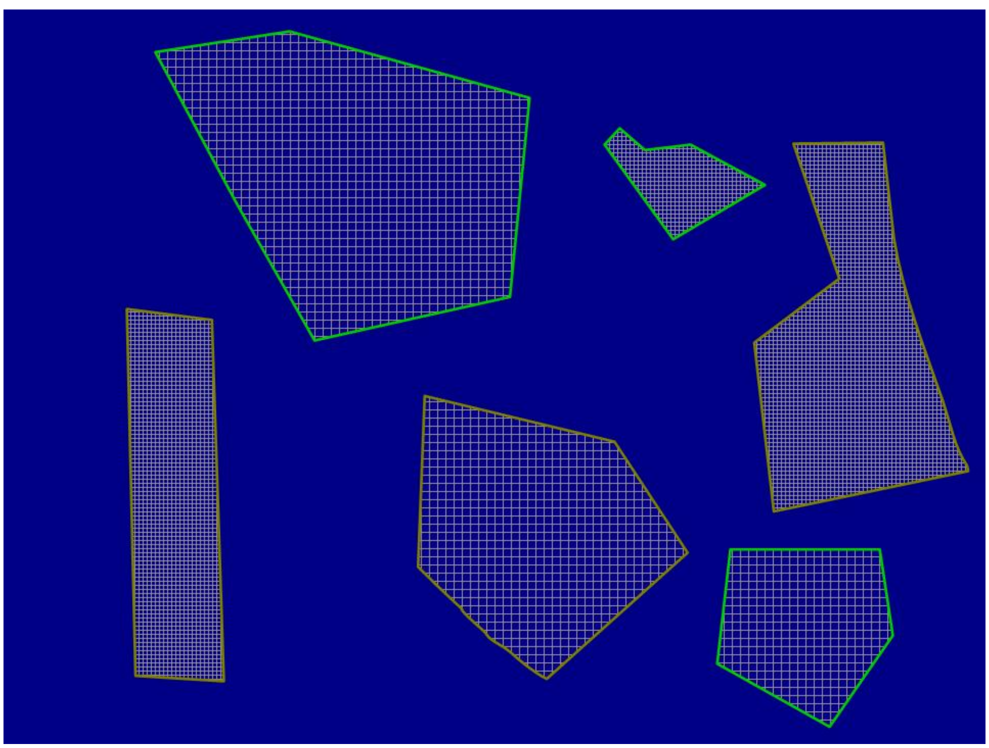
# Task 3

- Please use your mouse to move the cursor over and point out the coverage regions that have WEAK coverage quality



Integrating People, Performance & Technology™

24



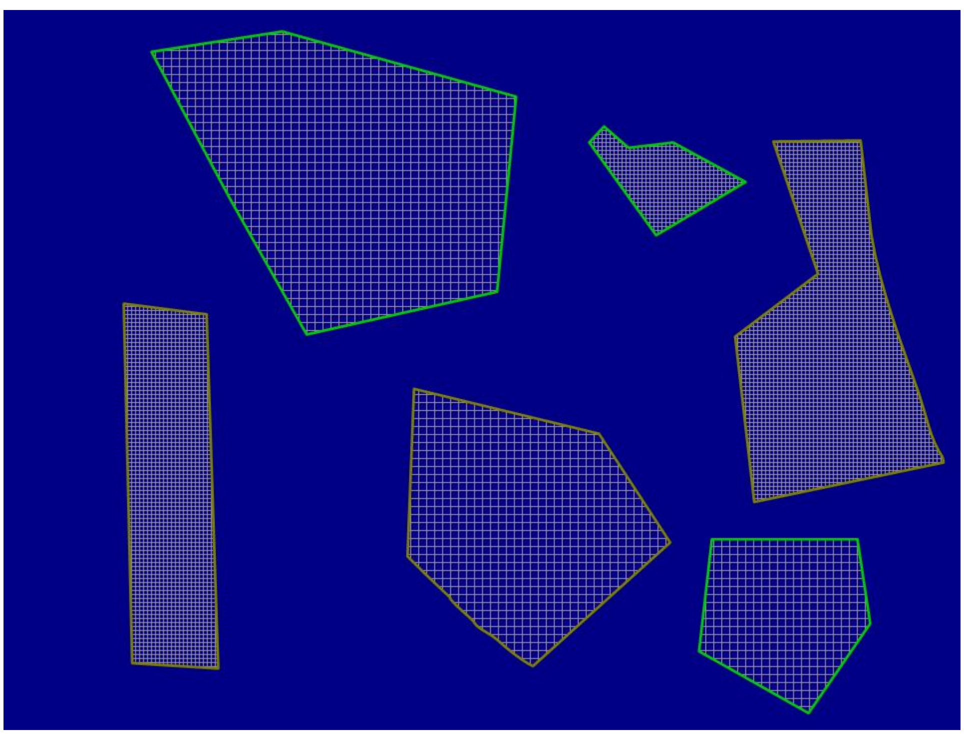
# Task 4

- Please use your mouse to move the cursor over and point out the coverage regions with:
- ***strong*** QUALITY INFORMATION and,
- ***out of date*** TIME INFORMATION

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

26



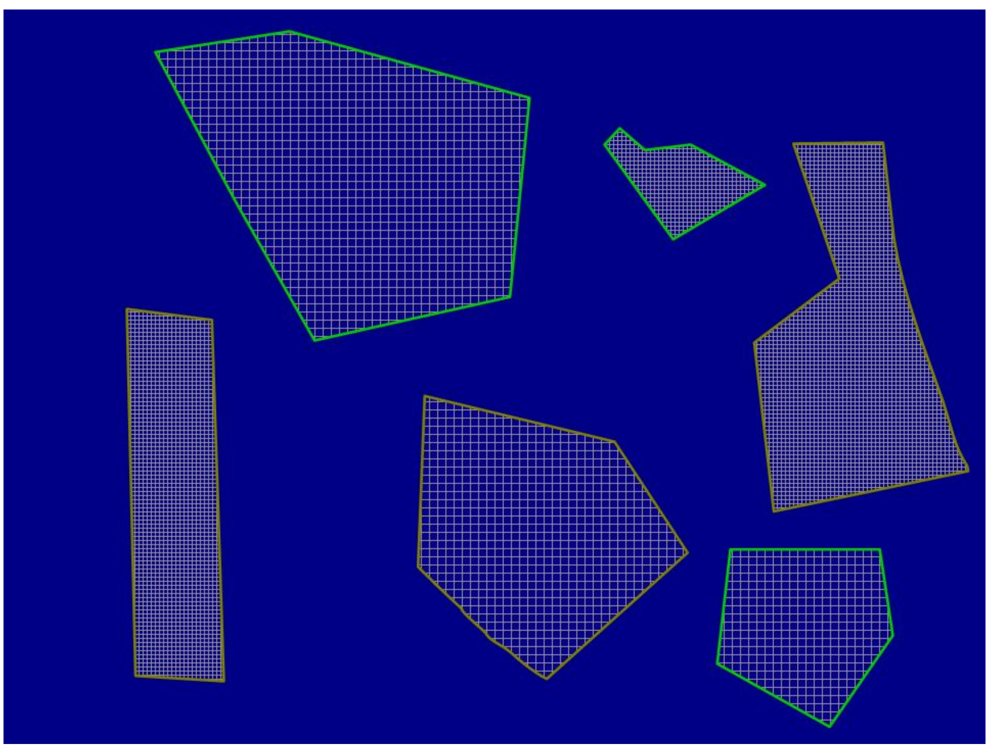
# Task 5

- Move on to the next slide and find all the coverage regions with:
- ***weak*** QUALITY INFORMATION and,
- ***current*** TIME INFORMATION



Integrating People, Performance & Technology™

28



# Task 6

Now things are going to get more complicated!

Move on to the next slide and find all contacts in **coverage** regions with:

- ***strong*** QUALITY INFORMATION and,
- ***current*** TIME INFORMATION

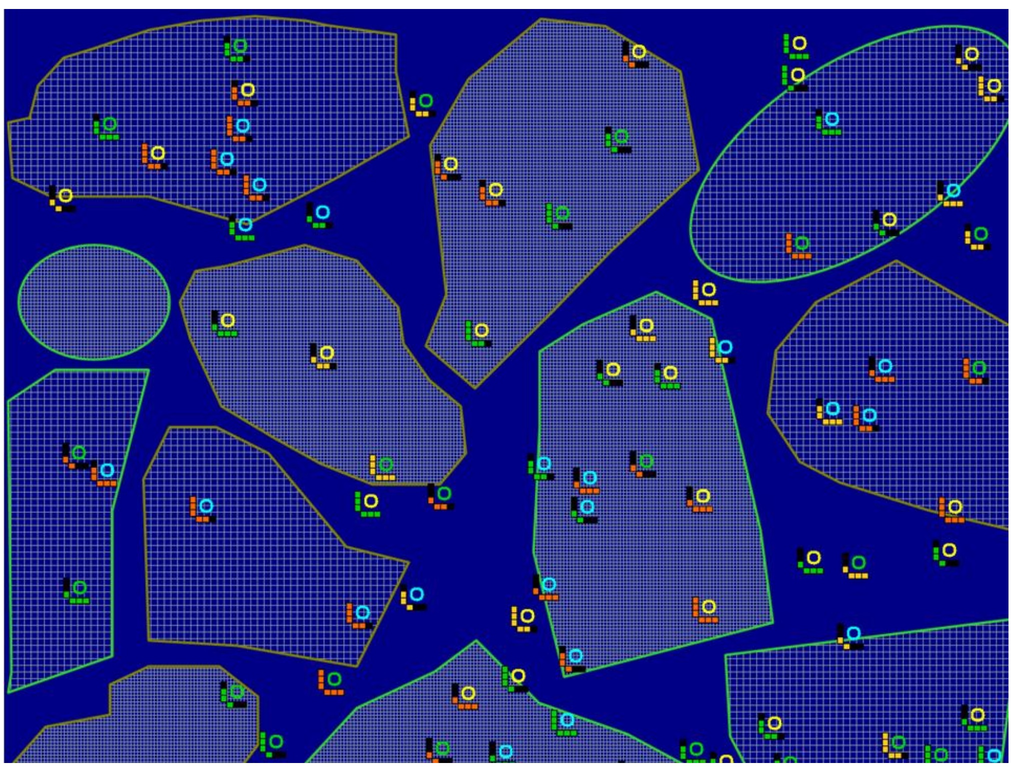
AND where the **contact** symbols are **completely within** the border lines having:

- ***current*** time information
- ***unknown*** spatial position information
- ***no*** identity information

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

30



# Task 7

Move on to the next slide and find all contacts in **coverage** regions with:

- ***weak*** QUALITY INFORMATION and,
- ***out of date*** TIME INFORMATION

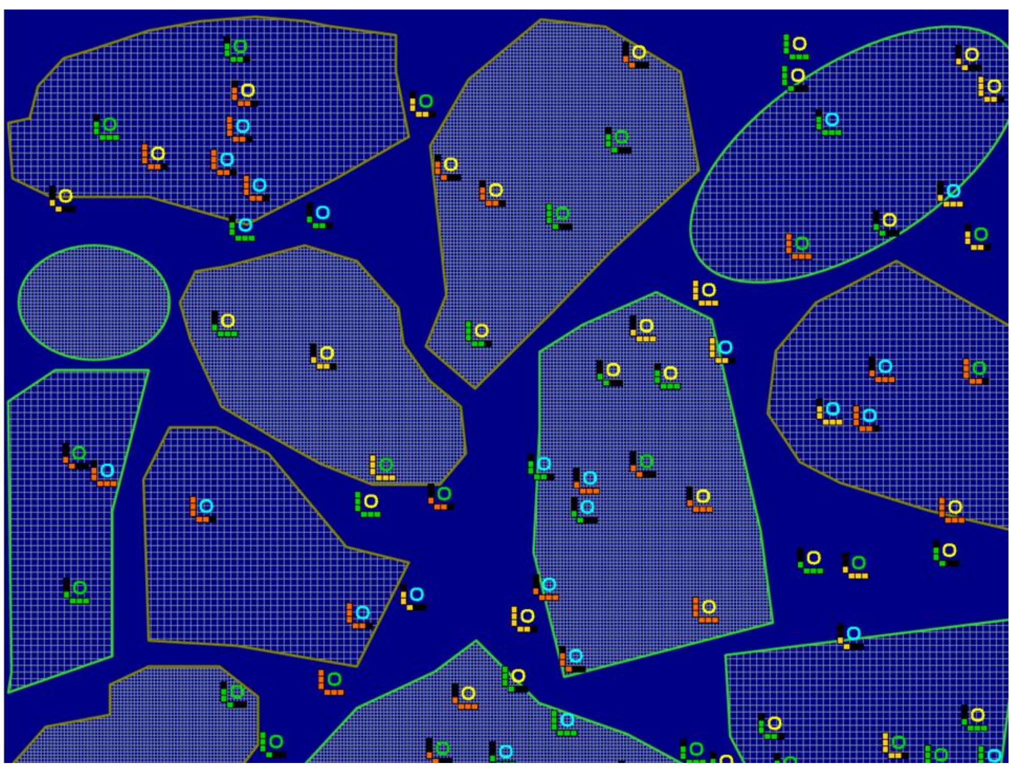
AND where the **contact** symbols are **completely within** the border lines having:

- ***out of date*** time information
- ***approximate*** spatial position information
- ***no*** identity information



Integrating People, Performance & Technology™

32



# Task 8

Move on to the next slide and find all contacts in **coverage** regions with:

- ***weak*** QUALITY INFORMATION and,
- ***current*** TIME INFORMATION

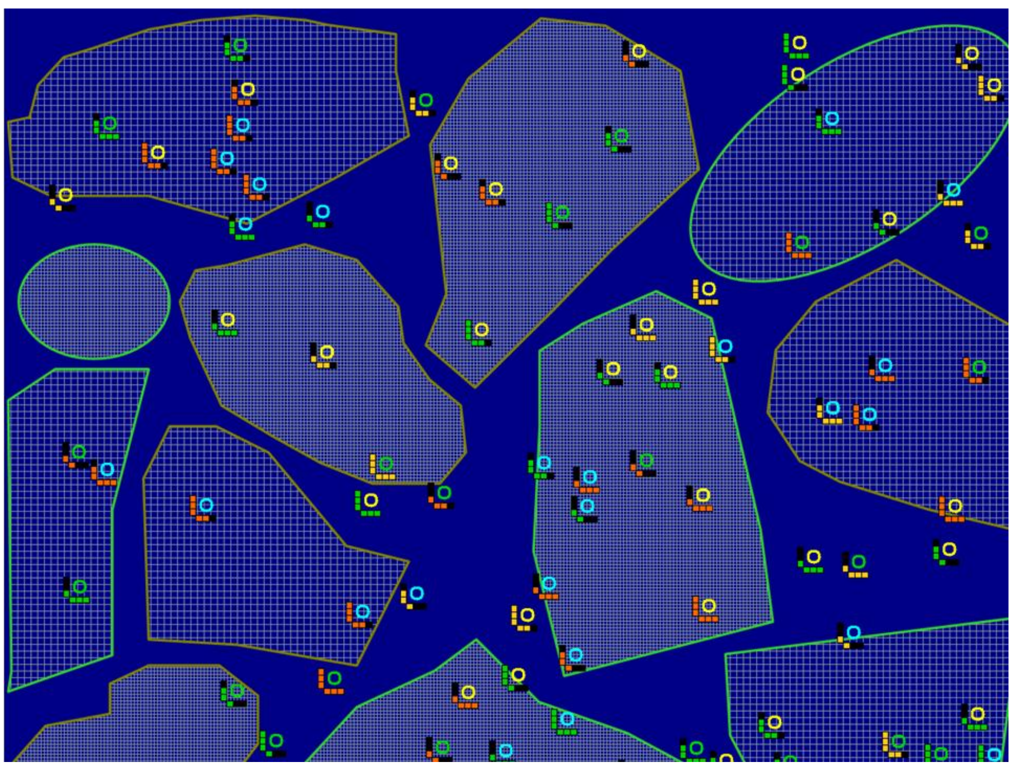
AND where the **contact** symbols are **completely within** the border lines having:

- ***current*** time information
- ***precise*** spatial position information
- ***good*** identity information



Integrating People, Performance & Technology™

34



# Task 9

Move on to the next slide and find all contacts in **coverage** regions with:

- ***strong*** QUALITY INFORMATION and,
- ***out of date*** TIME INFORMATION

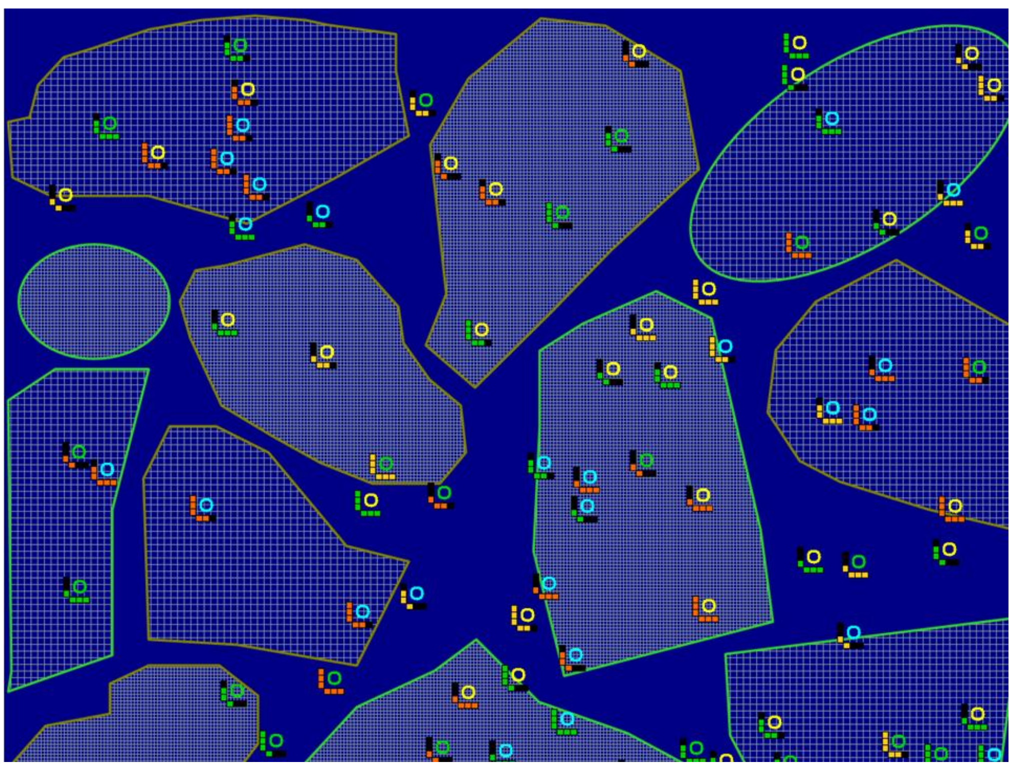
AND where the **contact** symbols are **completely within** the border lines having:

- ***out of date*** time information
- ***precise*** spatial position information
- ***good*** identity information

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

36



# Contact information

- **Human*Systems*® Incorporated**
  111 Farquhar Street,
  Guelph, Ontario N1H 3N4
  1 519 836 5911
  www.humansys.com

- **Michael Matthews**
  1 519 836 5911
  mmatthews@humansys.com

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

38

# Thank you!

- Thanks again for participating in this study and for completing the training in preparation for the study.
- If you have any questions please feel free to ask the experimenter who will be happy to answer them for you.

©Humansystems® Incorporated
www.humansys.com

Integrating People, Performance & Technology™

39

# Distribution List

Document No.: DRDC Atlantic CR 2008-177

| Copies | LIST PART 1: Internal Distribution by Centre |
|---|---|
| 5 | DRDC Atlantic LIBRARY |
| 1 | M. MCINTYRE |
| 1 | T. HAMMOND |
| 1 | A.W. ISENOR |
| 1 | S. WEBB |
| 1 | A. MACINNIS |
| 1 | SECTION HEAD OF MICS |
| 1 | J. CREBOLDER |
| 1 | B. CHALMERS |
| 1 | D. COADY |
| 1 | M. HAZEN |
| 1 | L. LAPINSKI |
| 16 | TOTAL LIST PART 1 |

| | **LIST PART 2: External Distribution by DRDKIM** |
|---|---|
| 1 | Library and Archives Canada |
| 1 | DRDKIM |
| 3 | DRDC Valcartier<br>2459 Pie-XI Blvd North, Québec, Quebec G3J 1X5<br>Attn: Alain Auger<br>Dennis Gouin<br>Alain Bouchard |
| 1 | DRDC Toronto<br>1133 Sheppard Avenue West, PO Box 2000, Toronto, Ontario M3M 3B9<br>Attn: Sharon McFadden |
| 1 | DRDC Ottawa<br>3701 Carling Avenue, Ottawa, Ontario K1A 0Z4<br>Attn: Chris Helleur |
| 3 | Human*systems*® Incorporated<br>111 Farquhar St., 2nd floor Guelph, Ontario N1H 3N4<br>Attn: Michael Matthews<br>Lisa Rehak<br>Ron Boothby |
| 1 | Commanding Officer<br>TRINITY JOSIC, PO Box 99000 Stn Forces, Halifax, NS B3K 5X5 |
| 1 | Capt(N) K. Greenwood<br>J3, Canada COM HQ, NDHQ - 101 Col. By Dr., Ottawa, ON K1A 0K2 |

1 LCdr A. Gyorkos
OIC Maritime Operations Centre, JTFP HQ, MARPAC/JTFP,
PO Box 17000 Station Forces, Victoria, BC B9A 7N2

13 TOTAL LIST PART 2

**29 TOTAL COPIES REQUIRED**

**DOCUMENT CONTROL DATA**
(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

| 1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's document, or tasking agency, are entered in section 8.) | 2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) |
|---|---|
| Publishing: DRDC Atlantic<br>Performing: Humansystems® Incorporated, 111 Farquhar St., 2nd floor, Guelph, ON, N1H 3N4<br>Monitoring: DRDC Atlantic<br>Contracting: DRDC Atlantic | UNCLASSIFIED |

3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title)

Evaluation of New Visualization Approaches for Representing Uncertainty in the Recognized Maritime Picture (U)
Évaluation de nouvelles techniques de visualisation permettant de représenter l'incertitude dans la situation maritime générale (U)

4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.)

Michael Matthews, Lisa Rehak, Julie Famewo, Tamsen Taylor, Jeremy Robson

| 5. DATE OF PUBLICATION (Month and year of publication of document.) | 6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) | 6b. NO. OF REFS (Total cited in document.) |
|---|---|---|
| October 2008 | 116 | 12 |

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Contract Report

8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development − include address.)

Sponsoring:
Tasking:

| 9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.) | 9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.) |
|---|---|
| 11he | W7701−054996/002/QCL; Call−up No. W7707−078067 |

| 10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document) | 10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.) |
|---|---|
| DRDC Atlantic CR 2008−177 | |

11. DOCUMENT AVAILABILITY (Any limitations on the dissemination of the document, other than those imposed by security classification.)

Unlimited distribution

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited announcement

# UNCLASSIFIED

**DOCUMENT CONTROL DATA**

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)


(U) This report documents the literature review and experimentation used to develop and assess visualization options to represent uncertainty in the Recognized Maritime Picture (RMP), which is the visual representation of the surface vessel picture for the Canadian maritime Area of Interest (AOI). Specifically, visualization options for the uncertainty with regards to the identity, spatial position and time lateness of surface contacts and the quality and time lateness of the sensor coverage were developed and assessed using computer−based experiments at the Humansystems (HSI„µ) Test Lab. Two icons (Rectangle design and ¡§Lego¡¨ design) were developed to display uncertainty related to the surface contacts, in addition to background swaths with two features (fill and border) to display sensor coverage uncertainty. Search times and accuracy were explored through 6 experimentation sessions with 11 participants. The results showed a small search time advantage for the Rectangle design and small performance differences among the different designs for sensor coverage. Participants rated the workload associated with using the designs as low. All of the design options evaluated are considered to be suitable candidates for future evaluation by the operational community. This work was conducted as part of the Information Visualization and Management of Enhanced Domain Awareness in Maritime Security Applied Research Project within the Defence Research and Development Canada (DRDC) Maritime Domain Awareness (MDA) research thrust.

(U) Ce rapport rend compte de l’analyse documentaire et de l’expérimentation qui ont servi à élaborer et à évaluer des moyens visuels de représenter l’incertitude de la situation maritime générale (RMP), c’est−à−dire la représentation visuelle des navires de surface qui empruntent la zone d’intérêt maritime du Canada. Plus précisément, on a élaboré et évalué des moyens de représenter visuellement l’incertitude en ce qui concerne l’identité, l’emplacement et la tardivité des contacts de surface et la qualité et la tardivité de la couverture des capteurs, dans le cadre d’expériences informatisées réalisées au laboratoire d’essai de Humansystems (HSI). On a conçu deux icônes (le modèle Rectangle et le modèle Lego) pour illustrer l’incertitude relative aux contacts de surface, en plus de champs de fond (bandes) comportant deux attributs (remplissage et bordure) permettant d’indiquer l’incertitude de la couverture des capteurs. Le temps de recherche et l’exactitude des données ont été examinés lors de six séances d’expérimentation qui ont réuni 11 participants. Les résultats montrent que le temps de recherche est légèrement réduit lorsqu’on utilise le modèle Rectangle et que le rendement des différents modèles applicables à la couverture des capteurs est variable. Les participants ont évalué comme faible la charge de travail associée à l’utilisation des modèles. Ils estiment que tous les modèles conceptuels examinés méritent d’être évalués plus en profondeur par des utilisateurs. Ces travaux ont été réalisés dans le contexte du projet de recherche appliquée intitulé Visualisation et gestion de l’information pour accroître la vigilance dans le secteur maritime, qui relève du vecteur de Recherche sur la connaissance du secteur maritime (CSM) de Recherche et développement pour la défense Canada (RDDC).


14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)


(U) Information Visualization; Uncertainty Visualization; Uncertainty; Visualization; Recognized Maritime Picture; RMP

This page intentionally left blank.

**Defence R&D Canada**

Canada's leader in defence
and National Security
Science and Technology

**R & D pour la défense Canada**

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE R&D DÉFENSE

**www.drdc-rddc.gc.ca**