Citation:
Katre Dinesh (2004). Experimenting with UniFace: A tool for Usability Testing of Icons, ACM-SIGCHI-SI National Usability Conference Easy 2004, Bangalore, India.

UniFace: Internet based Software for Remote Usability Testing of Icons

Dinesh S. Katre
Centre for Development of Advanced Computing (C-DAC)
Ministry of Communications and Information Technology, Government of India,
Agriculture College Campus, Near District Industries Centre,
Shivaji Nagar, Pune 411005, India.
Tel/Fax: 91+020+5533250
Personal Homepage: www.hceye.org

ABSTRACT
The Graphical User Interface (GUI) of software usually consists of huge number of icons. Though the intention is to improve the usability of software, not all interface designers are able to test and evaluate the comprehensibility of icons. Increasing exposure to unevaluated icons causes cognitive fatigue to users and slows down the intuitive learning. Users from diverse geographic locations, cultures and religions are very likely to interpret and understand these icons differently. As software products are designed to address universal needs, testing and evaluation of GUI across the globe or at least, wherever the product is likely to be used becomes important. Creation of dedicated usability labs in various locations for usability testing is not a viable proposition. A software tool named '*UniFace*' for remote usability testing of icons is designed capitalizing on far-reaching capability of Internet. *UniFace* extends the usability lab onto the desktop of every user. It encourages stakeholder participation in the design process and captures their perceptions. The testing methods of *UniFace* produce various reports with measurable and eloquent data for empirical analysis. The database of icon properties and test results can be helpful in visualizing the cognitive models of various user groups. *UniFace* has the potential to facilitate the standardization of iconic language for GUI design. It can prove very useful for offshore software projects.

Keywords
GUI, Icons, Remote Usability Testing, Cognitive Model, Measurable Data, Stakeholder Participation

FLOOD OF ICONS

The sign language of icons is becoming complex day-by-day due to increasing number of icons introduced through operating systems, software packages and websites. Its vocabulary is ever evolving and multiplying very rapidly. Our observations reveal that a novice computer user needs familiarity with minimum 40 to 50 icons on Windows platform for basic word processing using MS Word. An average computer professional has familiarity with at least 300 to 400 icons. When counted the number of icons on a computer of a multimedia professional, it crossed the thousand mark. The users get exposed to more and more icons as they navigate through various software products. Young computer users can take extra efforts for learning new icons and memorize many of them but the older people show significant decline in their ability to recall [13]. Resultantly, the training needs have grown out of shape, as many novice users need to learn the definitions of icons. Organizations have to spend a lot of money since users take paid time to get trained [28]. This proves the fact that not all iconic interfaces are comprehensible and usable as expected.

Need of Measuring the Usability of Icons

Most designers tend to test the comprehensibility of icons on their teammates or on a few subjects or simply ignore testing by documenting the definitions of icons as part of online help. The excessive and unevaluated usage of such icons must be causing significant cognitive stress, irritation and prolonged learning for software users. The testing reveals that several icons are actually misleading, ambiguous, conflicting and unnecessary [25, 15, 11]. In the past, there have been major research explorations for identifying the guidelines of interface designing [7]. On the basis of these guidelines, most interface designers build their hypothesis of how users would interpret the icons. But how effectively the designer has succeeded in transferring the theoretical guidelines into design [30] needs to be measured and evaluated based on factual data [21].

Constraining Factors of Usability Testing

The sign language of icons has to be universalized [19] for precise communication especially when software products are addressing universal applications and utilities. Intercultural issues referring to the religious, historical, linguistic, aesthetic, and other more humanistic issues of particular groups or peoples, sometimes crossing national boundaries need to be addressed for achieving greater acceptability to the products [20]. The generality and adaptability [2] of iconic interfaces can be achieved through externalization of our thoughts, ideas and concepts [8] for inviting criticism and user participation in the design process [11]. But the prohibitive factor is the cost of setting up usability labs for testing of iconic interfaces in geographically distributed locations [27]. Involving usability experts and renting the facilities for such testing makes it an unrealistic proposition. All such impeding factors highlight the importance of "discount usability engineering" approach [22]. In this situation, the only option is to explore remote usability testing methods [23].

In summary, the observations and requirements of usability testing narrated above necessitate the development of remote usability testing methods and a far-reaching mechanism for encouraging participation of users from diverse geographic locations.

Difficulties of Interface Design Students & Instructors

C-DAC's National Multimedia Resource Centre, Pune, India conducts training programs to teach multimedia. The students are introduced to various interface design issues with special emphasis on interactive multimedia software. In an exclusive academic project, the students are asked to design icons representing various links and functions of software. Every year, around 84 students produce approx. 1500 icons for varied themes of software. Evaluation of these icons is extremely challenging and a mammoth task. Awarding higher grades to icons, which are very easily understandable but not very well rendered, often invites disagreement of students. Many times beautifully rendered icons, yet based on ambiguous metaphors [5] do not communicate well. The instructor grades the projects based on his/her personal understanding of icons. The difference of viewpoints between the students and the instructor results into disputes at the time of grading. This is due to unavailability of definite testing methods and inability to measure / quantify the usability of huge number of icons designed by students.

The frequent disagreements with students about grading of their iconic interfaces forced us to devise a mechanism that could enable them in exploring third party evaluation themselves. The outcome of this experiment is found beneficial to not only the students but also the usability experts.

By and large these are common problems faced by most interface designers and usability experts. Testing methods like open-ended written comprehension, multiple choices have been already proposed and experimented for finding out the plausible response to a set of symbols [10, 25, 3, 6, 31]. As documented in Wolff's ANSI report, a large hall with proper seating arrangement, printed booklet with the symbols for testing and two judges were required for conducting the tests. Major amount of time was invested for collating the feedback of subjects. The Internet based solution described in this paper incorporates existing as well new methods for confirming the results.

UNIFACE: AN INTERNET BASED METHOD FOR USABILITY TESTING & EVALUATION OF ICONS

UniFace offers very simple, clear and result oriented Internet based usability testing methods. Eloquence of data captured by *UniFace* reduces the dependency on usability experts [16]. Following are the prominent features / characteristics of *UniFace*-

- Tests the comprehensibility / usability of icons within the design phase.
- Allows conducting of remote usability tests in diverse geographic locations for wider user participation in design process.
- Captures and maintains the record of user profiles and test results.
- Stores the database of icons with variety of attributes.
- Helps interface designers in visualizing the cognitive models of defined user groups.
- Produces measurable statistics and usable reports for substantiating the decisions [21].
- Provides a simplified procedural template for usability testing of icons that can be used by small-scale software developers.

Table 1. Basic Modules of *UniFace*

Sr. No.	Names of Modules	Performed by	Activities
I	**Integration**	Interface Designer	To integrate the icons along with their properties
II	**Testing**	Targeted Users (Subjects)	To record the feedback
III	**Evaluation**	Usability Expert	To observe, interpret and judge the feedback of subjects

Introduction to the basic modules of *UniFace*

UniFace is developed using web technologies and has three modules namely, Integration, Testing and Evaluation as shown in Table 1. Access to users and designers is controllable project as well as module wise.

Some aspects relating to 'Attributes of Icons' mentioned in this paper are based on information already reported in journals. This information is briefly given in the paper so as to fully communicate the testing methods and the design of *UniFace*.

Module I- Integration of Icons

Having designed the icons, the designer has to start with the Integration Module of *UniFace*. One begins by creating a separate project with appropriate name. A project can include various segments / subsets of icons. Each segment can be named appropriately as e.g. toolbar, editor, etc. Figure 1. illustrates the overall structure of Integration Module.

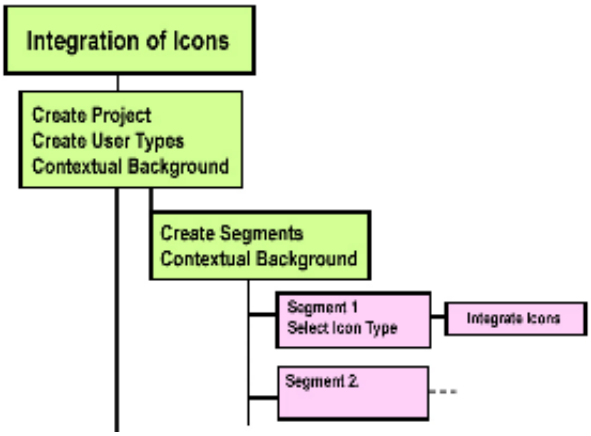


Figure 1. Design of Integration Module

Every icon has certain inherent attributes such as rendering style, type, form, associated proposition, title and visual elements. The interface designer has to indicate the attributes of an icon by selecting appropriate options at the time of integration. *UniFace* maintains a record of these attributes in the database. The definitions of the basic attributes of icons are elaborated hereafter.

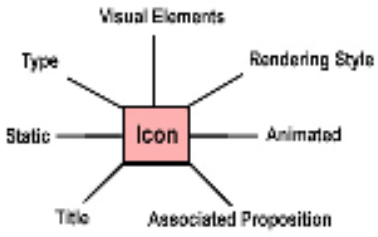


Figure 2. Attributes of an Icon

The designer has to indicate the pictorial quality of icons in terms of 'optimized' and 'moderate' rendering styles [16]. Brief explanation of these rendering styles is given below-

Optimized Rendering

Many icons used in software packages such as MS Office are rendered in optimized manner. The icons are pretty small in size (16 x 16 or 32 x 32 pixels). These icons are oversimplified and have just minimum necessary details.

Moderate Rendering

If compared with optimally rendered icons, moderately rendered icons are quite detailed and colorful. The icon size is much larger (64 x 64 pixels and above).

Rendering styles have to be indicated, as the recognition of optimized icons is at times slower than that of moderately rendered icons. Chances of a person recognizing the moderately rendered icon are higher. Also, the rendering style has an implication on recall value of an icon.

In addition, the designer has to select whether the icons belong to object, action and concept types [29]. Following are the definitions and critical aspects relating to each type of icon.

Object Icons

As these icons represent real life objects they are very easy to illustrate and recognize. The weakness of these icons is lack of proportional details, which at times leads to misinterpretation by user.

Figure 3. Object Icons

Action Icons

'Action Icons' are very difficult to illustrate and understand as depiction of an action or a process involves two or more stages. The designer often uses arrows while depicting the element of time or before and after stages in action icons. Action icons can be best communicated in the form of an animation.

Figure 4. Action Icons

Concept Icons

The designer tries to build analogies by showing some arbitrary [25] or metaphoric images while representing abstract features of software. Users from different cultures are very likely to misinterpret the concept icons.

Figure 5. Concept Icons

Form

Animated icons are found very effective in places where static icons fail to deliver the message. The software products have both forms of icons (static and animated) used as part of the GUI.

Associated Proposition and Title

Associated proposition means the intended message to be communicated by an icon. This message is then compressed in one or two words as the title of an icon. Users usually remember the icon with its title if both are resonating the same meaning.

Visual Elements of an Icon

Visual elements used in icons such as computer, arrow, brush, globe and chain, tick mark, etc. also have to be mentioned as part of icon attributes in *UniFace*. We come across varied permutations and combinations of these visual elements composed for communicating different shades of messages.

For the first time, *UniFace* has integrated and woven all attributes of icons together as important parameters of usability testing and analysis.

Define Stratified User Groups

Having integrated all icons as mentioned earlier, the interface designer has to now define the targeted user categories. These are broad categories of users such as tourists, accountants, medical practitioners, teachers, students, etc. It is expected that users representing all such categories be tested through *UniFace*. If a particular category of users is not tested then the report produced by *UniFace* indicates the same.

UniFace supports creation of user accounts under every stratified group by entering their basic information like name, company, e-mail address, etc. Projects already integrated in *UniFace* can be allotted to users for which they comply the eligibility criteria. Unique User IDs and passwords get e-mailed to users along with a list of projects.

Contextual Background

The designer has to provide a small write-up about the software product in *UniFace,* as recognition and interpretation of icons is quite context dependent. In addition, it is possible to link the screenshot of overall interface layout showing other related icons. The subjects can read the write-up and view the screenshot of interface layout before starting the test.

Integration of Multiple Clues

At this stage, the designer is supposed to input total three clues for every icon. It consists of one clue expressing the desired proposition and the other two clues communicating the probable and yet misleading propositions. Having completed this preparatory work, *UniFace* gets ready for performing various tests.

The batch of Sept. 2002 to March 2003 of Diploma in Advanced Computer Arts course tested around 13 GUI projects using *UniFace* at C-DAC, Pune, India. The students accessed *UniFace* on Intranet and were permitted to perform the tests on minimum 6 subjects per project due to time constraints. Overall around 74 subjects participated in the tests. Broad characteristics of test results are derived based on observations recorded during the experiments.

MODULE II- TESTING

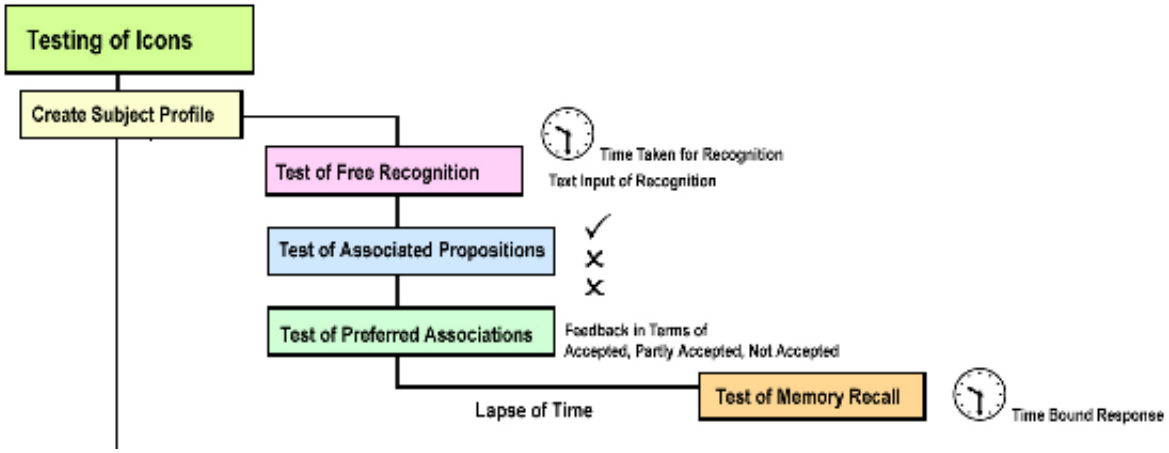


Figure 6. Design of Testing Module

Objectives of Testing

Interface designers and usability experts must be very clear about the objectives of testing. Various tests as indicated in Figure 6. are performed to find out whether the icons are-

-easy to interpret and understand.
-communicating precise message.
-easy to memorize and recall.
-ambiguous and confusing.
-are slow communicators.
-needing initial introduction.

Above objectives largely determine whether the icons pass or fail the test. One can capitalize on the feedback captured by *UniFace* for meeting the following objectives-

- Observe the trends and unique characteristics of response received from various user groups and improvise the icons.
- Try to visualize the mental model of subjects on the basis of data captured by *UniFace*.
- In case of favorable results, certify the icons for integration in the software.

Precautions

The tests performed using *UniFace* can produce best results if certain precautions are taken. They are as under-

- Send an e-mail communication mentioning the User ID and URL for accessing *UniFace* and guidelines of the test to various user groups. Ensure that the subject profiles are properly matching with the groups defined earlier.
- Inform the subjects that the intention is not to examine their abilities but to get their feedback for improvising the GUI.
- The subjects must go through the test alone and not in a group. They should refrain from showing the testing process to others. If their colleagues or friends are interested in participating in the testing process then they must do it individually.
- The subject should be requested to fill their profiles sincerely and completely.
- While performing the test they should avoid other interruptions and complete the test without spending time anywhere else.

Create Subject Profile

It is mandatory for the subject to enter his profile details in the beginning of the testing module. Figure 7. shows the template for entering the subject profiles. It requires information on items such as Name, Age, Gender, Computer Awareness, Academic, Economic, Cultural / Ethnic background, Geographic Location and User Group. This information gets stored in the database of *UniFace*.

After completely filling the required profile details, the Subject is then allowed to begin with the first 'Test of Free Recognition'.

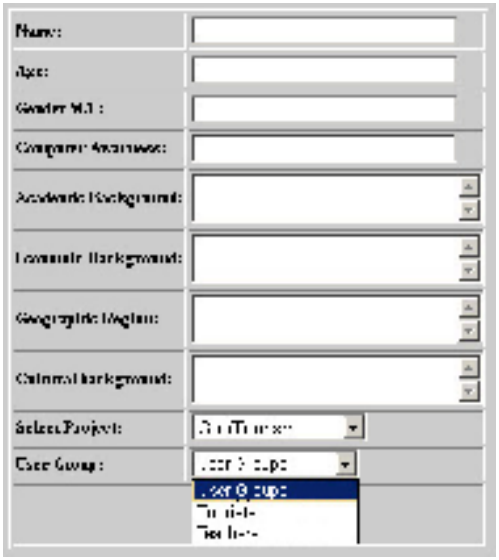


Figure 7. Template for Preparing Subject Profile

Test of Free Recognition (TFR)

In this particular test, only one icon is displayed on computer screen at a time without mentioning its title or associated proposition or clues as shown in Figure 8. The subject can wait until he has understood the icon and then type his response in the input area. In case of failing to recognize the icon, the subject can mention the same and move on to next icon. *UniFace* calculates and records the time taken for recognition starting from the display of icon up to clicking on 'next icon' link in the input area. The subject can refer the contextual information if needed.

Figure 8. Test of Free Recognition (TFR)

The TFR generates feedback about whether the visual elements rendered in the icon are recognizable and the visual semantics is conveying the associated proposition. The interface designer immediately comes to know whether the keywords are closer to the intended message of the icon or else. If the icons are designed with semiotic qualities then the rate of recognition is higher and faster as the test progresses. If the first icon is understood then the forthcoming icons are also grasped quickly but inconsistent and conflicting representations cause confusion [29].

The response generated by TFR can be characterized as under-

- The record of time indicates that ambiguous icons require longer time for recognition.
- The subject-
  - fails to understand an icon even after taking long time.
  - ends up describing the visual elements depicted in an icon but fails to capture the associated proposition.
  - associates an icon with some other function of software or misinterprets.
  - correctly associates an icon with the desired function or utility of software in short time.

Ambiguous icons indicate a common trend of long time taken for recognition. Subjects attempt blind guesses as well. The recognition input apparently includes a lot of irrelevant information but it is very useful. Subjects are often emotional and very elaborate in their response. You come to know about their vocabulary, proficiency of language and conversance with computer keyboard. Novice computer users make a lot of typographic errors.

The Test of Associated Propositions (TAP) becomes active as soon as the subject is through with TFR.

Test of Associated Propositions (TAP)

TAP displays every icon along with three clues, which are entered by the interface designer during the Integration Module of *UniFace*. The scope of interpretation is reduced to three plausible choices.

Icons help us in building a mental association between the visual and the feature of software. If the association between an icon and the desired proposition is not meaningful then the user is often misled. TAP captures the probability of an icon misleading the subject.

The response generated by TAP can be characterized as under-

- Many times, the icon, which has failed during TFR but is within the proximity of its message, gets recognized during TAP.
- TAP is useful in identifying icons that need initial introduction. If an icon and its associated proposition have semantic connection then after initial introduction, the users are able to understand it effortlessly.
- Some icons failed during TFR continue to evoke wrong responses.
- The icons that were properly recognized in the TFR are confirmed during TAP.

The user response reveals impreciseness of an icon e.g. Figure 9. shows that an icon of 'binocular' is designed to represent 'site-seeing' for a website on tourism. Many subjects selected 'binocular' as its correct recognition. This reveals that the element of 'site-seeing' is missing in the icon. Here, the icon has succeeded in communicating the referent but failed to communicate the proposition associated with the referent [15].

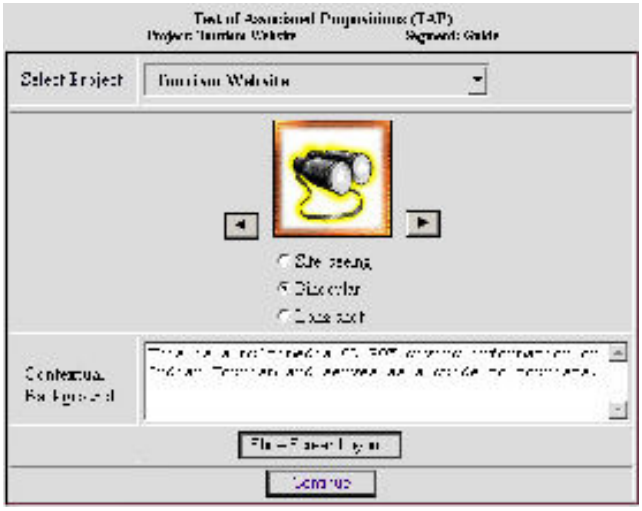


Figure 9. Test of Associated Propositions (TAP)

One may do away with such imprecise icon but what would happen if the standard 'find' feature is required in the website? As the typical 'find' feature is already represented by 'binocular' icon only. You come across many such loosely designed icons in software.

Test of Preferred Associations (TPA)

During this test, an icon along with its associated proposition is displayed on screen. The Subject is requested to give his verdict in terms of whether an icon is accepted, not accepted or partly accepted. The usability expert can corroborate the verdict with the actual response captured during TFR and TAP. TPA produces quantifiable data.

The response generated by TPA can be characterized as under-

- Icons that have failed during TFR and passed during TAP are unanimously accepted.

- Icons that have failed during TFR and TAP unanimously rejected. The subjects very often advise the possible alternatives.
- Subjects suggest improvements / modifications while partly accepting the icon.

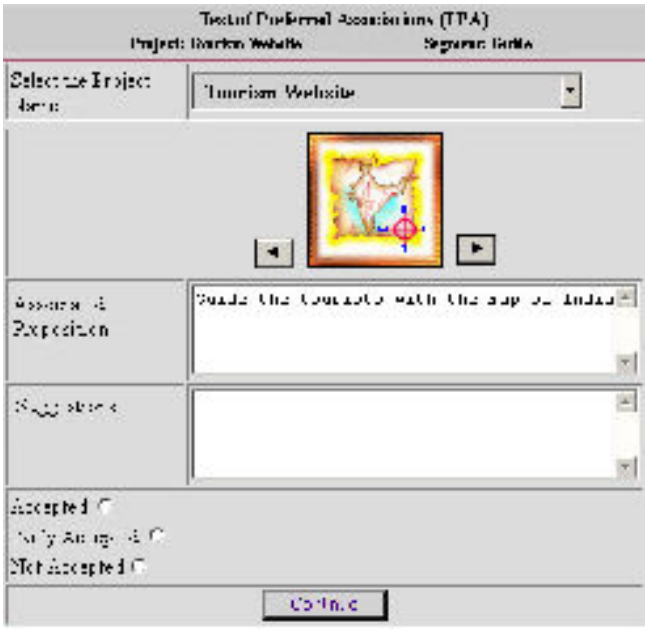


Figure 10. Test of Preferred Associations (TPA)

Test of Memory Recall (TMR)

This test is performed after certain lapse of time to find out whether the subjects are able to recall the icons along with their titles. The lapse duration can be based on frequency of software usage. TMR proves that meaningful association between visual and the title is recalled effortlessly.

So far, in all the tests (TFR, TAP, TPA) only one icon was displayed at a time. TMR displays maximum five icons simultaneously as the cognitive limit of an average person allows tracking of maximum seven items (plus or minus two) [1, 21]. On initiating the test, the identification titles of icons begin to appear on the screen one after another. The subject is expected to click the matching icon within stipulated duration before the next title appears on screen. Having matched all the icons, the subject can go for next lot of icons.

The subject is not given unlimited time for searching or remembering the icon. The usability expert can decide the display duration for every title on the basis of usability goals of GUI design. If the display duration is 5 seconds then the titles randomly change after every 5 seconds.

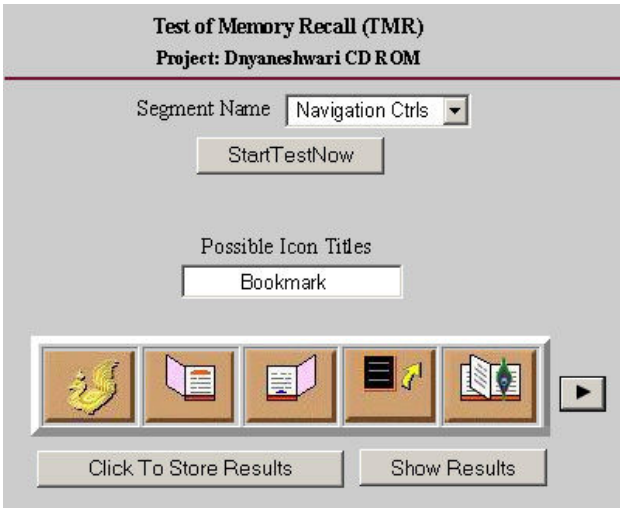


Figure 11. Test of Memory Recall (TMR)

The response generated by TMR can be characterized as under-

- The subject fails to associate the proposition with an ambiguous icon.
- The subject is confused with resembling icons and hence clicks on a wrong icon or fails to make a decision within stipulated time. During the testing of a tourism website, the subjects got confused between the icons of 'water sports' and 'beaches'. This reveals another aspect of impreciseness of an iconic expression.
- Obvious icons are correctly matched with their titles.

TMR can be very useful for testing the recalling ability of older people.

MODULE III- EVALUATION

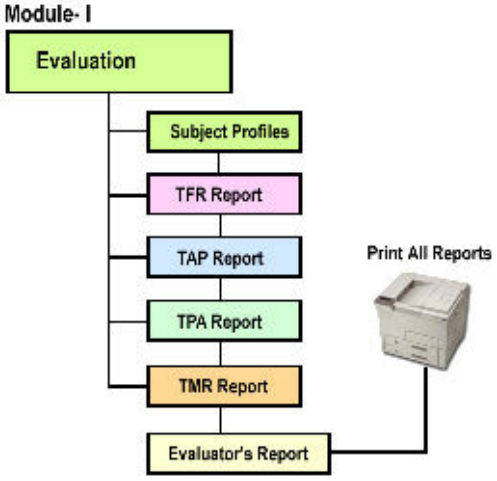


Figure 12. Design of Evaluation Module

Diagnostic Report

Icon-Type: obj
Icon Id:- 3

Expected Result : Lighthouse
User Types:
1 . Domestic Tourists
2 . Foreign Tourists

Lapse Duration: 4 Days

User Names	UserType	TFR Report	TFR Time Record	TAP Report	TPA Report	TPA Report Suggestions	TMR Report
srinivasa naidu	Domestic tourists	not recognisalbe/icon to b big	65	Lighthouse	Partly accepted	could have been zoomed in to show the details of the main light	Lighthouse
saumitra rath	Domestic tourists	light house	40	Lighthouse	not accepted	use sub dued method for this one representating the light house	Lighthouse
Manas Palo	Domestic tourists	not identifiable	35	Tower	not accepted	visual is not at all clear	Lighthouse
Sandeep	Domestic tourists	Not identifiable	26	Lighthouse	accepted	No comments	Lighthouse

nexticon previcon

Evaluation Assesment and Recommendations:-	
Accepted:	1
Not Accepted:	2
Partly Accepted:	1
Average time taken:	41.5

Acceptance (yes/no) No

If Acceptance with modifications,give details below

The visual is very ambiguous. The TFR time record shows that user is taking long time for recognition. TPrA response of 3 users is not favorable.

Figure 13. Diagnostic Report

The consolidated diagnostic report as shown in Figure 13. presents overall view of test results pertaining to an icon. The quantifyable aspects of TFR and TPA are presented at the bottom of the report where the usability expert can record his/her decision. Only 4 test results are shown in Figure 13. in order to fit the screenshot in smaller size.

The interface designer has to consider the feedback given by subjects and redesign the faulty icons again. The redesigned icons have to be tested and evaluated through *UniFace* through the same process until there is sufficient acceptance from subjects.

Printouts of all the reports along with the subject profiles are taken and attached to the final diagnostic report. The Interface designer and usability expert have to sign the final report.

A typical evaluation report of a tourism website finalized by the usability expert is shown in Figure 15. In this report, the usability expert has cleared 4 icons without any modification, 7 icons with minor modifications and 3 icons are rejected.

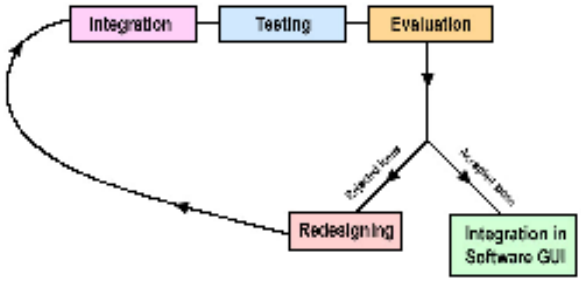


Figure 14. Cycle of Evaluation

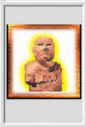
IconId	Icon Preview	Title	Modifications	Accepted	Not Accepted	Partly Accepted	Evaluator's Decision
tour-01		Beach	NA	5	1	0	yes
tour-02		Metro Cities	Ambiguity of visual may be reduced.	2	2	2	No
tour-03		Close	The closed door aspect should be more clear. Other decorative details should not standout.	4	1	0	Yes
tour-04		Open	NA	5	0	0	yes
tour-05		Restaurant	NA	4	1	0	yes
tour-06		Forts	More clarity in the visual form.	4	0	1	yes
tour-07		History	may be used with some improvization for depcting scriptures and not for History.	3	2	0	No
tour-08		Map Guide	NA	5	0	0	Yes
tour-09		Museums	Museum aspect should be depicted as this icon may conflict with Harappan Gallery Icon.	4	0	1	Yes
tour-10		Gardens	May be used for conveying greetings.	1	3	1	No
tour-11		Siteseeing	Siteseeing element should be added.	4	0	1	yes
tour-12		Transport	Road transport should also be represented.	5	0	0	Yes

Figure 15. Final Evaluation Report

The cycle of evaluation uncovers certain functions or features that are very difficult to represent in the iconic form. Redesigning and exploration of various visual metaphors do not produce positive results, as not all messages can be communicated effectively using representations [24]. In such cases, the Usability experts may decide not to use any iconic representations and instead use a simple button or a link with its title. This approach can be useful in reducing the pollution of incomprehensible and unnecessary use of icons.

The testing process of *UniFace* reveals the items that should not be shown as icons or the risk associated with it.

Iconographic imagery helps define a larger concept, [4] the symmetry between 'conceptual model' of software and 'interface metaphor'. Indirectly, *UniFace* also collects feedback about the underpinning interface metaphor manifested through icons. If majority of the icons successfully pass through all the tests then the comprehensibility of overall metaphor is also very high.

UniFace is introduced as part of the training programs conducted at C-DAC's National Multimedia Resource Center. The students are able to integrate their interface design projects and test them using *UniFace* on their own. The feedback gathered from the subjects is now agreeable to students without any difference of opinion. The grading of interface design projects is based on empirical analysis. Thus, *UniFace* can completely eliminate the aspect of 'gut feeling' involved in icon

designing. One can ensure that icons are really comprehensible enough to the targeted users.

DATABASE OF ICON ATTRIBUTES

Responses captured during TFR and suggestions given in TPA reports can be corroborated with subject profiles for figuring out the mental model of a user group. Of course inputs received during the test may not be adequate for defining the mental model. Usability expert will have to pick up the threads from test results and investigate the subject further.

Table 2. shows how the attributes of icons are stored in the database of *UniFace*. The database is helpful in evolving a lexicon of icons and avoiding overlapping usage of visual elements. There is an icon of Bookmark in Table 2. that shows an open book with a peacock feather as bookmark. This icon is designed for a multimedia CD on *Bhagavadgita* (Hindu religious book). *Bhagavadgita* was first recited by Lord *Krishna* (Hindu God). As per Indian mythology, Lord *Krishna* always has a peacock feather in his crown and therefore we chose to show peacock feather as a bookmark for *Bhagavadgita* CD ROM. This connotation of the icon is specific to Indian culture. This icon need introduction in other regions that India. Thus *UniFace* can capture the problems pertaining to cultural interpretations of icons through its testing process.

Text based search can be performed for identifying the icons that include similar visual elements. Repository of test results and database of icons can be used to find out the following-

- What types of messages are already communicated using similar visual elements?
- Whether the use of proposed visual elements is overlapping with other icons or is it unique enough?
- Do we already have some icon test results of subjects with required profile? This can be a useful input to another project.
- Have those visual elements (icons) positively cleared the tests?
- How recently the tests were performed?
- If subject profiles are suitable then what type of depiction of proposed visual elements they are already familiar with?

The data stored by *UniFace* can be used for prejudging the icons and guiding the efforts of user interface designing in right direction. Variety of testing methods, database of icons with attributes, collection of test results, and accessibility through Internet shows the potential of *UniFace* to facilitate the standardization of iconic language.

CONTINGENCIES AND REMEDIAL OPTIONS

With *UniFace*, it is not possible to monitor the respondents, as they operate from remote locations. In an uncontrollable testing mode, there are chances of a few subjects not performing the tests sincerely. Failing to understand the procedure is also likely if the subjects ignore the guidelines and precautions. But we don't escape these occupational hazards even while testing in a usability lab. For countering such problems, the usability expert can maintain a good dialogue with the subjects through e-mail communication.

Table 2. Properties of icons stored by *UniFace*

Icon	Title	Associated Proposition	Type	Rendering Style	Form	Visual Elements
	Annotate	Attach annotated comments	Object	Optimized	Static	Paper, Pen
	Status	Indicate status in terms of chapter, pages, verses	Concept	Optimized	Static	Dial, Needles
	Date and Time	Inform the date and time	Concept	Moderate	Static	Calendar, Clock, Dial, Needles
	First Page	Jump to first page	Object	Optimized	Static	Open Book, First Page, Cover
	Help	Open help documents	Concept	Moderate	Static	Closed Book, Question Mark
	MS Word Help	Provide necessary help	Concept	Optimized	Static	Balloon, Question Mark
	Settings and Help	Provide necessary help and controls for settings	Concept	Optimized	Static	Question Mark
	Bookmark	Go to bookmark	Object	Optimized	Static	Open Book, Peacock Feather

Being the stakeholders of proposed software product, the subjects should be convinced about importance of testing and their feedback for positively influencing the design process.

Adequate testing experience of *UniFace* can help the interface designers / usability experts in judging the sincerity of a subject. In case of doubt, one may just entirely disqualify the feedback. Correspondingly, it necessitates that sufficient number of subjects (minimum 20/25 subjects per user group) must be tested.

Even though *UniFace* is far reaching, it is still confined to the parts of world where Internet has spread and sufficient bandwidth is available. Network of cyber-cafes and regional distributors of software products could be involved where the users are unable to access Internet. Involving true stakeholders in the testing process can produce better quality of feedback [9].

At present, only English is supported as part of *UniFace*. Support of regional languages may be required for achieving better communication with subjects from diverse regions.

CONCLUSIONS

UniFace has retained the positive aspects of conventional usability testing methods for evaluation of icons and overcome the constrains such as-

1. Dedicated usability labs are not required. It expands the scope of testing without escalating the cost.
2. It records the time taken by every individual user for TFR.
3. Automates the random shuffling between the titles and icons for TMR.
4. It captures and collates the test results in the form of reports and presents statistical data for analysis.
5. Crosses the geographical boundaries through Internet and allows wider user participation. This can help in revealing the cross-cultural aspects connected with the interpretation of icons.

In addition, *UniFace* provides a reliable testing framework for improving the quality and precision of communication through icons. It offers following advantages.

6. Usability expert can take decisions substantiated by enough evidence and measurable data captured by *UniFace*. The reports generated by *UniFace* clearly indicate which icons have passed or failed the tests and suggestions for improvement.
7. Software companies can invite participation from various user groups / stakeholders through *UniFace* while designing the GUI.
8. Testing through *UniFace* can be extremely helpful in ensuring that the GUI becomes a true cognitive mediator [18]. Cognitive stress on users and learning time can definitely be reduced to a certain extent.

ACKNOWLEDGEMENTS

I am grateful to Dr. Aaron Marcus, President, AM & A, USA for his encouraging feedback and supportiveness. I am thankful to Prof. Jan Gulliksen, Director of Studies (Human Computer Interaction), Dept. of Information Technology, Uppsala University, Sweden for his valuable suggestions. I acknowledge the valuable guidance of Dr. Mukul K. Sinha, my Ph.D. Supervisor and Managing Director of Expert Software Consultants, Pvt. Ltd. and the support of my group members at C-DAC.

REFERENCES

1. Baddeley, A. 'Your Memory' A Users Guide, Published by PRION, UK, (1996), 78-79.
2. Beaudouin-Lafon, M. User interface support for the integration of software tools: an iconic model of interaction, in Proceedings of the third ACM SIGSOFT/SIGPLAN software engineering symposium on Practical software development environments, (Boston, Massachusetts, United States, 1989), 143- 152.
3. Baecker, R., Small, I., and Mander R. Bringing icons to life, Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology (New Orleans, Louisiana, United States, March 1991), 1-6.
4. Byrne, D. M. Using Icons to Find Documents: Simplicity Is Critical, In the Proceedings of INTERCHI, (Amsterdam, The Netherlands, April 1993), 446-453.
5. Brami, R. Icons: A Unique Form of Painting, ACM Interactions, (Sept. + Oct. 1997), 15-28.
6. Britton, Jr. D., and Reyes, A. Discovering Usability Improvements for Mosaic: Application of the Contextual Inquiry Technique with an Expert User, The Second International WWW Conference '94: Mosaic and the Web Electronic Proceedings, 1994. Available at http://archive.ncsa.uiuc.edu/SDG/IT94/Proceedings/HCI/britton/britton_reyes.html
7. Carroll, M. J. Introduction: Human Computer Interaction, the Past and the Present, Human Computer Interaction in the New Millennium, Published by ACM Press and Addison-Wesley, 2002.
8. Fischer, G. Beyond Couch Potatoes: From Consumers to Designers, Third Asian Pacific Computer & Human Interaction Conference Proceedings, (Kanagawa, Japan, July 1998), 15-17.
9. Giordano, R. and Bell, D. Participant Stakeholder Evaluation as a Design Process, in Conference

Proceedings of ACM Universal Usability, (2000), 53-60.
10. Hemenway, K. Psychological issues in the use of icons in command menus, In the Conference Proceedings of the SIGCHI conference on Human factors in computing systems, (Gaithersburg, Maryland, United States, 1982), 20-23.
11. Holtzblatt, K., and Beyer, H. Apprenticing with the Customer: A collaborative Approach to Requirements Definition, in Electronic Proceedings of Communications of the ACM, (May 1995).
12. Holloway, J., and Bailey, J. Don't Use a Product's Developers for Icon Testing, in Proceedings of CHI '96, (Vancouver, British Columbia, Canada, April 1996), 309-310.
13. Howthorn, D. Possible Implications of Aging for Interface Designers, Interacting with Computers 12, Elsevier Science, (2000), 507-528.
14. Ivory, M., and Hearst, M. The state of the art in automating usability evaluation of user interfaces, ACM Computing Surveys (CSUR), Volume 33, Issue 4 (December 2001), 470-516.
15. Katherine, H. Why icons cannot stand alone, ACM SIGDOC Asterisk Journal of Computer Documentation, Volume 20, Issue 2 (May 1996), 1-8.
16. Katre, D., Unconventional Inspirations for Creating Software Interface Metaphors, in Proceedings of International Conference on Media and Design (Mumbai, India, Sept. 2002), Volume I, 1-15.
17. Lackoff, G., Metaphors We Live By, Published by Cambridge Press, (1990).
18. Laurel, B. Computers as Theatre, Published by Addison-Wesley, 1993, 7th Edition in 1999.
19. Marcus, A., Icons, symbols, and signs: visible languages to facilitate communication, ACM Interactions, Volume 10 , Issue 3 (May + June 2003), 37 - 43
20. Marcus, A., Armitage, J., and Frank, V. Globalization of User-Interface Design for the Web, 5th Conference on Human Factors and the Web, in Electronic Proceedings, (1999) Available at http://zing.ncsl.nist.gov/hfweb/proceedings/marcus/index.html
21. Miller, G.A., The magical number seven, plus or minus two: Some limits on our capacity for processing information, The Psychological Review, (1956), 81-97.
22. Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D., and Kirakowski, J. Comparative evaluation of usability tests, in Proceedings of the Conference on Human Factors and Computing Systems, (CHI 1999), 83 – 84.
23. Nielsen, J. Why GUI panic is good panic, Interactions, Volume 1, Issue 2 (April 1994), 55 - 58.
24. Nielsen, J., Offshore Usability, Alertbox, Useit.com, Septmember 16, 2002, Available at http://www.useit.com/alertbox/20020916.html
25. Poovaiah, R., Graphics Symbols for Environmental Signage: A Design Perspective, in A design Perspective, IDC Publication, March 1997, IIT Bombay, 29-42.
26. Potosnak, K. (Editor), Do icons make user interfaces easier to use?, Human Factors, (May 1988), 97-99.
27. Redish, J., Bias, R., Bailey, R., Molich, R., Dumas, J., and Spool, J. , Conference on Human Factors and Computing Systems, (2002), 885 - 890.
28. Rowley, E., Usability testing in the field: bringing the laboratory to the user, Conference proceedings on Human factors in computing systems, 1994, 252 – 257.
29. Schaffer, E., and Sorflaten, J. ICONS: Much Ado About Something, Human Factors International, UI Design Update e-Newsletter, (December 1998), Available at http://www.humanfactors.com/library/icons.asp
30. Shneiderman, B., and Mahajan, R. Visual and Textual Consistency Checking Tools for Graphical User Interfaces, IEEE Transactions of Software Engineering, Vol. 23, No. 11, (November 1997), 722-735.
31. Sutcliffe, A. On the effective use and reuse of HCI knowledge, ACM Transactions on Computer-Human Interaction (TOCHI), Volume 7, Issue 2 (June 2000), 197 – 221.
32. Wolff, J. A Study of the Effect of Context Aand Test Method in Evaluating Safety Symbols, Georgia Institute of Technology, Graphics Visualization and Usability Centre's Report No. GIT-GVU-96-07, 1996.