2009 13th International Conference Information Visualisation

The Correlation between the ability to Read and Manually Reproduce a 3D Image: some implications for 3D information visualisation.

Theodor Wyeld
Flinders University, Australia
{theodor.wyeld@flinders.edu.au}

Abstract

Most of us can recognize common 3D objects depicted in drawings, photographs and computer graphics. But, few of us are able to manually reproduce them in a convincing manner. This paper discusses a psychology experiment that investigates the variability in individual drawing ability and the ability to read 3D images. Currently, Shepard and Metzler's [1] Mental Rotation Test is the most popular test for spatial ability. This paper discusses the need to further investigate the correlation of 3D drawing ability and recognition and its potential effect on the legibility of the 3D information visualisation application. This paper reports ongoing research in this field.

Keywords--- **3D images, MRT, 3D information visualisation, chiaroscuro, perspective drawing**.

Introduction

Before humans could write we used pictures to communicate our thoughts. Cave drawings, petroglyphs and maps, are examples of picture languages. Primitive cave drawings are the earliest surviving picture language. They look a bit like children's drawings – flattened, with everything in view at once. When depicting objects in space, children often include different scales seen from different angles, all in the one picture. [2] describe this stage in a child's spatial development as egocentric. Rather than depicting just what can be seen, a child will depict what they know to be within their personal space. Their drawings include symbolic, metaphoric, or iconic elements, as well as realistic depictions [3]. Whilst we grow to understand an allocentric space, few of us ever develop our pictorial skills beyond the child's symbolic level. We might be able to recognise more sophisticated representations (such as photographs, 3D computer generated images and chiaroscuro drawings), but we often cannot manually reproduce or construct our own convincing 3D illustrations.

Where cave drawings present a flattened space, the ability to represent spatial depth is only a recent invention. In a rudimentary form, it first appeared in 1st century Byzantine, Chinese and Islamic art. But, it was not until the 15th century that a consistent method was formulated for constructing what we know today as the perspectival image [4]. Training in this new visualisation schema reached its zenith in the seventeenth century at the École des Beaux-Arts. Clearly, to manually produce a convincing 3D perspective image we need some training in its conventions. Today, a lot of the construction of perspective has been automated. 3D computer graphics programs now produce convincing perspective images with little or no understanding needed on the part of the producer in the algorithms used in their construction.

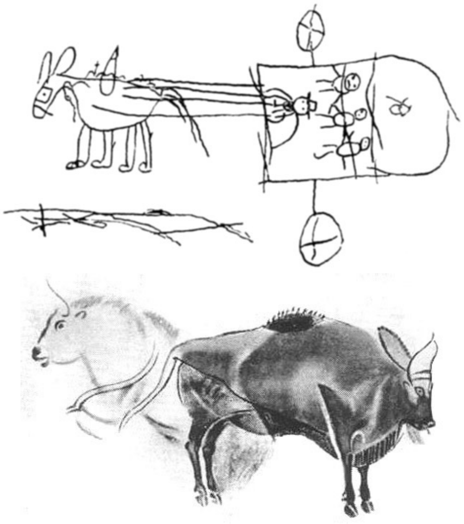

Figure 1 Child's drawing of horse and cart and 30,000BC Spanish cave drawing [5, p69 & 97].

In the field of information visualisation the 3D perspective image is used to create abstract organised representations of complex data sets. The added depth dimension helps us visualise the 'space' of the data. These abstract 'information spaces' are mapped onto 3D geometry with the added dimensions of: colour, texture, transparency, animation, sound, haptic feedback and so on. There are many convincing information visualisations that use the added depth dimension present in a 3D display [see 6, 7, 8, 9]. These pictorial displays can provide an abstract overview of metadata while preserving its intrinsic properties. The algorithms used to create these computer-generated perspectives are based on the same perspective construction techniques developed in the sixteenth-century Italian Renaissance.

978-0-7695-3733-7/09 $25.00 © 2009 Crown Copyright
DOI 10.1109/IV.2009.79

IEEE computer society

Much of the visualised information in a 3D format capitalises on the human capacity to see patterns in data collections. Patterns help us identify interrelationships. Compared to 2D, 3D representation adds an information dimension along the 3rd axis that follows the notion that data occupies space and that we might learn more if we could navigate that space. However, some researchers are reporting that presenting information in a 3D format does not always achieve the desired results – communicating a better understanding of the information being presented.

For example, [10] results on the use of 2D or 3D icons used in aircraft landing navigation favoured the vertical axis only; [11] 2D and 3D visualisation of telecommunications traffic found that 3D interfaces affected user performance in a negative manner; [12] comparative study of 2D and 3D targets in military studies identified 2D as more accurate; [13] two and a half dimensional format proved more efficacious than a full 3D display; [14, p271] claims that a "full-fledged 3D configuration could be too complex for users to handle efficiently", suggesting a 2.1D or 2.5D interface as a more useful compromise solution; [15, p57] report that users find it "more difficult to select objects in a 3D space compared to 2D"; and, [16] finding that full 3D views are not functional in their treegraph structures.

The inability for some users to efficiently recognise the 3D spaces depicted has led to alternative visual representation methods. For example, [13] argue that, what they call, a two and a half dimension format is less ambiguous in its representation of the information space. It relies on occlusion, scaling and ground plane alone. In another example, [17] uses a 3D network graph without object perspective cues (such as smooth, rounded, forms with specular highlights etc). Instead, it uses depth cues such as fading in the distance to reduce the overall sense of a volumetric space. This forces the user to focus on detail in the foreground. And, [16] adds only minor 3D cues to their, otherwise 2D, treegraphs in the form of cushioning, to indicate a textured surface thus reducing its abstractness and making it appear more organic, and perceptually more legible. These and other reductions on the 3D theme respond to the notion that not all users can efficiently read a full 3D display.

Despite the move to a reduced 3D thematic, none of these researchers attempt to address directly the reason *why* many of their users are more efficient at 'reading' the reduced displays. In the field of psychology there are many tests to determine one's spatial ability.

The most common test used to determine one's ability to recognise 3D objects and spatial arrangements is [1] Mental Rotation Test (MRT). In the MRT participants view a misaligned pair of 3D armatures constructed of cubes and decide whether they are identical (see figure 2). This test has established that most people can perform rudimentary mental transformations of objects. The amount of time it takes participants to perform the self-congruence task increases "monotonically with the angular disparity between objects… [suggesting] that individuals mentally rotate objects in the same manner in which they physically rotate objects" [18, p271]. Males are generally better in accuracy and speed in this test than females. The [19] version of this test uses different, more organic, objects (such as body parts) but returns similar results. Both tests require participants to recognise 3D object images on a screen or on a printed card to determine one's spatial ability. But, neither give us a clue to variability between individuals – why one participant might be better than another at the same task.

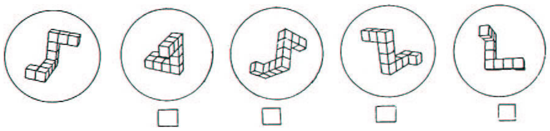

Figure 2 Original MRT test screens [20, p194].

In the standard MRT the only information available for identifying individual variance in spatial ability is the collected demographic data (gender, age, ethnicity). Unlike the MRT recognition test, if participants are required to draw, rather than simply identify, common 3D objects there is greater variability in ability than the MRT.

A person's ability to draw a common 3D object can be observed as their ability to accurately represent that object as a 3D artefact. In a group test, individual illustrations tend to fall across a range of abilities; from flattened, child-like to fully realistic chiaroscuro drawings.

This paper reports on an experiment conducted to determine what the range of variability in drawing a common 3D object is. The implications from the results suggests there may be a correlation with information visualisation researchers findings that a reduced 3D format tends to be more legible for some of their users.

The experiment described in this paper identifies variability across gender, ethnicity, and socio-economic status as it relates to the ability for a group of first-year university students to draw a convincing perspective of a common 3D object. The experiment was used to determine how a group of participants are able to self-rank their own drawings of a common 3D object from most to least realistic.

There were two parts to this experiment. In the first part, 50 first-year undergraduate university students were given 5 minutes to draw a box in perspective. In the second part, a different group of 55 first-year undergraduate university students were given 5 minutes to draw an apple in perspective. As a control, the second part addresses the notion that the box in the first part is an orthogonal object, and, as such, lends itself to depiction in a 3D perspective (receding lines etc) in a way that an apple does not. Nevertheless, both the box and the apple are common 3D objects that share some common perspective depth cues, such as shading, shadows and a ground plane.

Experiment

Participants were asked to indicate their gender, cultural background and socio-economic status by recording their gender, western or non-western cultural background, and attendance at private or public school. On each drawing participants made a marking of:

m = male, or f = female;
w = western, or n = non-western; and,
r = private school, or u = public school.

No other identifying marks were used. These classifications were used as opposed to others as they were deemed the least intrusive, thus, potentially generating the greatest compliance. Of the 105 students surveyed only 3 chose the option not to provide this information. As they represent such a small subset of the data collected their contribution is not included here.

The tables below show a breakdown of student participants by gender, cultural background, and socio-economic status for each part 01 (Box) and part 02 (Apple) of the experiment.

Table 1 Demographic data (Box)

Category	**G1**	**G2**	**G3**	**G4**	**G5**	**Tot**
Male (m)	6	12	2	7	8	35
Female (f)	2	2	4	5	0	13
Western (w)	8	12	3	7	4	34
non-Western (n)	0	2	3	5	4	14
Private (r)	2	10	5	4	5	26
Public (u)	6	4	1	8	3	22

Table 2 Demographic data (Apple)

Category	**G1**	**G2**	**G3**	**G4**	**Tot**
Male (m)	12	9	18	8	47
Female (f)	2	3	1	1	7
Western (w)	5	8	14	8	35
non-Western (n)	10	4	4	2	20
Private (r)	8	6	7	3	24
Public (u)	7	6	12	6	31

The tables below show the gender, cultural background and socio-economic status classifications as sub-groups for each super group.

Table 3 Demographic classifications (Box)

Classification	**G1**	**G2**	**G3**	**G4**	**G5**	**Tot**
mwr	2	8	3	1	2	16
mnr	-	-	-	1	3	4
mwu	4	2	1	-	2	9
mnu	-	2	3	-	1	6
fwr	-	2	1	2	-	5
fnr	-	-	-	1	-	1
fwu	2	-	2	-	-	4
fnu	-	-	2	1	-	3
Totals	**8**	**14**	**12**	**6**	**8**	**48**

Table 4 Demographic classifications (Apple)

Classification	**G1**	**G2**	**G3**	**G4**	**Tot**
mwr	3	5	5	2	15
mnr	5	-	4	1	10
mwu	1	2	9	4	16
mnu	3	2	-	1	6
fwr	-	1	-	-	1
fnr	-	-	-	-	-
fwu	1	1	-	1	3
fnu	1	1	1	-	3
Totals	**14**	**12**	**19**	**9**	**54**

Method

The same method was used for each part of the experiment (Box and Apple). Five groups of six to fourteen students participated in the first part of the experiment. Four groups of nine to nineteen students participated in the second part of the experiment. Participants were instructed to close their eyes and imagine a box with equal sides (or an apple in part 02). After approximately one minute they were asked to open their eyes. They were then asked to draw what they imagined. As a group, they were asked to self-rank all the drawings in order of the most realistic to the least realistic. Triangulation occurred where students ranked each other’s drawings describing and debating what constituted a ‘more realistic’ drawing, and why. In this way consensus was reached on why each drawing was lodged in its ranked order before moving on to the next one. Sometimes re-ranking occurred (moving of drawings into a position at the more or less realistic end of the array). This again required consensus before the final positions could be agreed upon.

Results

The results of the participants' drawings and their subsequent ranking was analysed statistically by calculating the spread of ranked values across all categories (m, f, w, n, r, and u). From the analysis and comments, there is no difference in a participant’s ability to produce a more or less realistic perspective depiction of a box regardless of gender, ethnicity or socio-economic status. However, in analysing their rationales for *why* they ranked their drawings in the ways they did participants reached consensus that they represented what they had been taught rather than what they had imagined in their mind’s eye. In other words, they did not draw what they saw in their mind’s eye but they replayed the actions they had been taught regarding how to draw a convincing perspective box (or apple in part 02) – they ‘visualised’ the finished drawing rather than the object itself. Their execution of this visualised drawing varied according to their individual ability. Those who’s drawings were ranked most realistic self-report more comprehensive training in depicting objects in perspective than those who’s drawings were ranked the least realistic. There was no other variability across all factors. This was confirmed by graphing and ANOVA analysis.

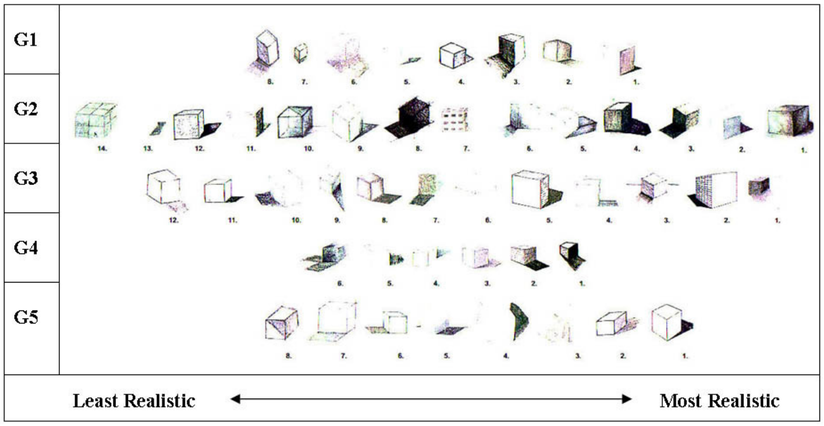

Figure 3 Actual box drawings by all participants by group showing consensual self-ranking order from most to least realistic.

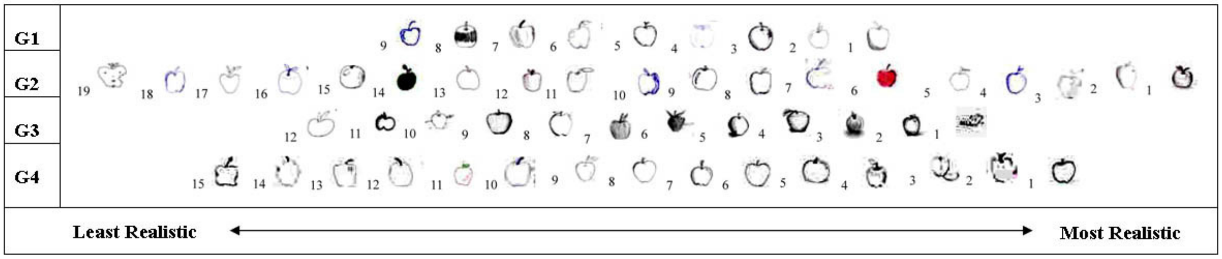

Figure 4 Actual apple drawings by all participants by group showing consensual self-ranking order from most to least realistic.

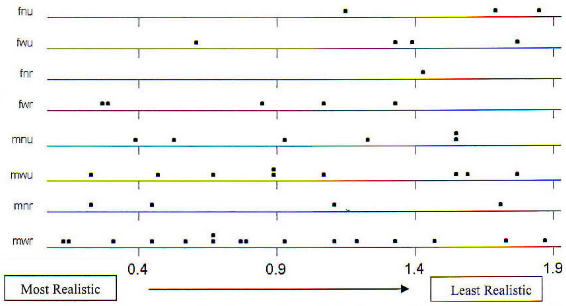

Figure 5 All participants in all classifications showing correlation between classifications related to relative position in averaged ranking order from most to least realistic (Box).

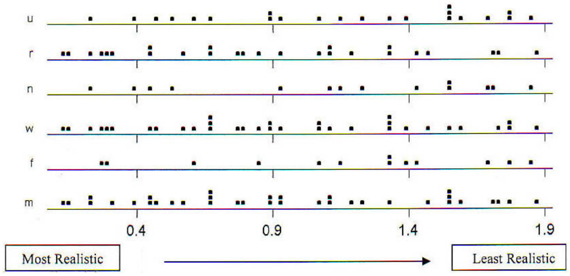

Figure 6 All participants in all categories showing correlation between categories related to relative position in averaged ranking order from most to least realistic (Box).

The graphs above demonstrate a mostly even spread of ranked positions, either by classification or category. Chi-square analysis further suggests no significant deviation from one category to the next. By combining all categories at once all classifications were compared using ANOVA analysis. This showed that the probability of any classified group, fnu versus fwu ...etc , producing a more or less realistic box was fairly weak (see tables below for Chi-Square and ANOVA Results).

Table 5 Chi-Square Tests (computed only for 2x2 table. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.28) (Box).

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.023[b]	1	.312		
Continuity Correction[a]	.516	1	.473		
Likelihood Ratio	1.027	1	.311		
Fisher's Exact Test				.385	.237
N of Valid Cases	47				

Table 6 ANOVA ranking (Box)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.576	5	.315	1.171	.327
Within Groups	37.155	138	.269		
Total	38.731	143			

The F statistic of 1.71 with 5 degrees of freedom only has a significance level of 0.327 (it should be 0.05 if there is a significant difference between the individual groups m, f, w, n, r, u), hence probability of a correlation is fairly weak.

A similar result was found in the apple test from graphing, Chi-Square and ANOVA analyses.

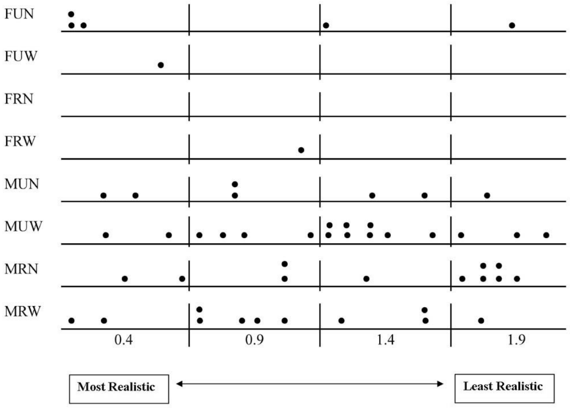

Figure 7 All participants in all classifications showing correlation between classifications related to relative position in averaged ranking order from most to least realistic (Apple).

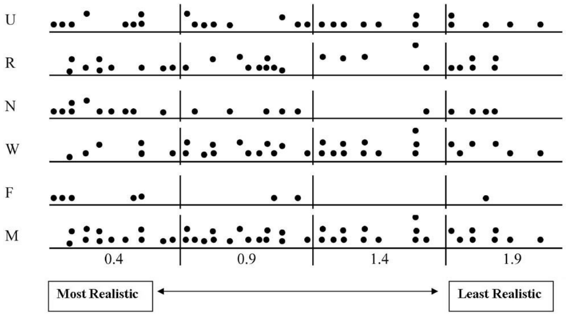

Figure 8 All participants in all categories showing correlation between categories related to relative position in averaged ranking order from most to least realistic (Apple).

Table 7 ANOVA ranking (Apple).

Source of Variation	Sum of Squares	d.f.	Mean Squares	F
between	1.375	5	0.2751	0.9083
error	48.15	159	0.3028	
total	49.52	164		

The probability of this result, assuming the null hypothesis, is 0.48. Assuming that the F statistic of 0.9083 with 5 degrees of freedom (it should be 0.05 if there is a significant difference between the individual groups m, f, w, n, r, u), hence probability of a correlation is fairly weak.

Table 8 Chi-Square Tests (Apple).

Male Group: Number of items = 47 Mean = 1.0752 95% confidence interval for Mean: 0.9166 thru 1.234 Standard Deviation = 0.528 High = 1.998 Low = 0.1540 Median = 1.100 Average Absolute Deviation from Median = 0.454	**Female Group: Number of items= 8** Mean = 0.68212 95% confidence interval for Mean: 0.2979 thru 1.066 Standard Deviation = 0.615 High = 1.846 Low = 0.1000 Median = 0.4810 Average Absolute Deviation from Median = 0.453
Westerner Group: Number of items = 35 Mean = 1.0747 95% confidence interval for Mean: 0.8910 thru 1.258 Standard Deviation = 0.492 High = 1.998 Low = 0.1540 Median = 1.100 Average Absolute Deviation from Median = 0.409	**Non-westerner Group: Number of items = 20** Mean = 0.91870 95% confidence interval for Mean: 0.6757 thru 1.162 Standard Deviation = 0.648 High = 1.875 Low = 0.1000 Median = 0.8220 Average Absolute Deviation from Median = 0.563
Private Group: Number of items = 26 Mean = 1.0062 95% confidence interval for Mean: 0.7931 thru 1.219 Standard Deviation = 0.565 High = 1.875 Low = 0.1540 Median = 0.9615 Average Absolute Deviation from Median = 0.462	**Public Group: Number of items = 29** Mean = 1.0320 95% confidence interval for Mean: 0.8302 thru 1.234 Standard Deviation = 0.552 High = 1.998 Low = 0.1000 Median = 1.110 Average Absolute Deviation from Median = 0.460

Discussion

Not all participants are equally able to produce a convincing perspective drawing. There was a lot of variability for both the box and the apple. However, the variability was uniform across all factors. All participants were able to recognise the cues that they and their peers used to describe their drawings as more or less realistic depictions of a common 3D object. Whether this reflects their ability to recognise other depictions of common 3D objects is not clear. From the drawings of the boxes it seems the conventions for representing a box in 3D are fairly well understood, albeit with variable skill. However, in the case of the apple there was greater variability in the representation methods. Some used shading and shadows whilst others appear to have used an outline only. This suggests that representing a box in 3D is better understood than an organic shape like an apple. It is not clear whether this is due to prior training, that the box better lends itself to perspectival representation than the apple, or that each simply represents the two major facets of a realistic drawing of a 3D object: perspective construction and chiaroscuro (combined they provide the most realistic impression). However, that illustrating a box seemed to be better understood, suggests this may go some way towards explaining why participants of the MRT have little trouble in recognising the objects used (orthogonal stacked box armatures). The implication from this is that people may need formal training in 3D image construction at an early age or we cannot expect them to be able to efficiently utilize them later in life.

Conclusion

Knowing how quickly and accurately people are able to recognise the orientation of similarly arranged armatures, constructed of cubes, floating in space (such as those in the MRT) suggests most people are able to recognise depictions of these types of 3D objects. However, when required to draw a 3D object in perspective results vary widely (more widely for organic than orthogonal objects). Hence, it seems recognition and representation may not be equivalent. The former requires experiential practice alone while the latter requires some training or practice. To-date, a test to correlate the relationship, if any, between one's ability to recognise a 3D object and the ability to draw it has not been conducted. In part, the experiment described in this paper addresses that gap.

The experiment conducted found that a participant's ability to draw a more or less realistic 3D object (a box or an apple) was statistically invariant regardless of gender, cultural background or socio-economic status. However, it also revealed that the participants in this experiment did not simply imagine a *real* box or apple before beginning their drawings. Rather, they visualised a *drawing* of a box or an apple. Furthermore, the drawing of a box or apple they visualised was based on how they had previously been trained to draw a box or an apple. Hence, their ability to draw a realistic box or an apple was contingent on how much training they had had prior to this experiment. It follows then that training in drawing common 3D objects may also improve one's ability to recognise other forms of 3D representation. This remains to be tested.

Those who reported having little training in drawing common 3D objects tended to produce the child-like flattened perspective drawings discussed at the beginning of this paper. In the context of information visualisation, as only a small percentage of participants were able to construct convincing 3D perspective drawings, this may go some way towards explaining why the 'flattened' perspective information visualisation display interfaces developed by Cockburn and McKenzie (2004) and others in response to some user preferences may be more legible than full 3D.

Perhaps with the demise of fundamental training in manual illustration skills in the depiction of common 3D objects in perspective – replaced by computer-generated perspective (3DCG, 3D Games, Photogrammetry and so on) – we may find fewer people able to reproduce what they can otherwise recognise. The fact that some information visualisation researchers are reporting greater legibility efficacies with reduced 3D displays suggests the replacement of manual skills with their computer-generated counterpart may not lead to improved 3D object recognition. On the contrary, we may find that as manual illustrative skills decline so too do people's ability to read 3D interfaces. The experiment described in this paper is provided as a precursor to a follow-up test to determine if indeed there is a correlation between manual 3D illustrative skill and spatio-visual abitlity to recognise the same objects produced artificially or by others manually.

Reference

[1] Shepard, R. N., and Metzler, J., 1971 Mental rotation of three-dimensional objects. *Science* 171: 701-703.

[2] Piaget, J. and Inhelder, B, 1956, *The Child's Conception of Space*, Langdon, F. J., and Lunzer, J. L. (trans.) Routledge and Kegan Paul, London.

[3] 1992 Taylor, H. A., & Tversky, B. (1992b). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31, 261-282.

[4] Hagen, M. A., 1986, *Varieties of Realism*, NY Cambridge University Press. (Chinese Persepctives).

[5] Kepes, G., 1967, *Language of Vision*, Paul Theobald, Chicago.

[6] Ware, C., 2004, *Information Visualization*, Morgan Kaufmann.

[7] Dodge, M., and Kitchin, R., 2002, *Atlas of Cyberspace*, Pearson Education.

[8] Chen, C., 1999, *Information Visualisation and Virtual Environments*, Springer-Verlag.

[9] Spence, R., 2001, *Information Visualisation*, ACM press, Addison Wesley.

[10] Wickens, C., C. Liang, T. Prevett and O. Olmos (1995). Egocentric and Exocentric Dipslays for Terminal Area Navigation, in *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*. 16--20.

[11] Hicks, M., C. O'Malley, S. Nichols and B. Anderson (2003). "Comparison of 2D and 3D Representations for Visualising Telecommunication Usage." *Behavior and Information Technology* 22(3): 185--201.

[12] Smallman, H., M. StJohn, H. Oonk and M. Cowen (2000). Track Recognition Using Two- Dimensional Symbols or Three-Dimensional Realistic Icons, *SPAWAR Systems Center*, San Diego.

[13] Cockburn, A., and McKenzie, B., 2004, "Evaluating Spatial Memory in Two and Three Dimensions", in the *International Journal of Human-Computer Studies*. 61(3): 359--373. Elsevier Science. 2004.

[14] Chen, C., 1998, "Bridging the Gap: The Use of Pathfinder Networks in Visual Navigation", in *Journal of Visual Languages and Computing* (1998) 9, pp267-286.

[15] Roth, S. F., Chuah, M. C., Kerpedjiev, S., Kolojejchick, J., Lucas, P., 1997, "Towards an Information Visualization Workspace: Combining Multiple Means of Expression" in *Human-Computer Interaction Journal*, Volume 12, Numbers 1 & 2, 1997, 131-185.

[16] Wijk, J.J. van, H. van de Wetering. "Cushion TreeMaps: Visualization of Hierarchical Information", in: G. Wills, D. Keim (eds.), Proceedings 1999 IEEE *Symposium on Information Visualization* (InfoVis'99), October 25-26, 1999, IEEE Computer Society, p. 73-78.

[17] Chen, C. (2005) Top 10 Unsolved Information Visualization Problems, in *IEEE Computer Graphics and Applications*, 25(4), 12-16. source: http://www.computer.org/portal/cms_docs_cga/cga/content/Promo/promo2.pdf

[18] Wraga, M., Swaby, M., and Flynn, C. M., 2008, "Passive Tactile Feedback Facilitates Mental Rotation of Handheld Objects", in *Memory & Cognition*, 2008, 36 (2), pp271-281.

[19] Vandenberg, S. G., & Kuse, A. R., 1978, "Mental rotations: A group test of three-dimensional spatial visualization", in *Perceptual and Motor Skills*, 47, 599–604.

[20] Shiina, K., Short, D. R., Miller, C. L, Suzuki, K., 2001, Development of Software to Record Solving Process of a Mental Rotations Test, in *Journal for Geometry and Graphics*, Volume 5 (2001), No. 2, pp193-202.