AN EXPERIMENTAL STUDY: TEXTUAL INFORMATION DRIVEN SPATIAL UNDERSTANDING AND REPRESENTATION FOR USER INTERFACE DESIGN OF 3D MODELING TOOLS

Li, Xuenan

School of Art Design and Media, East China University of Science and Technology

ABSTRACT

The same information has different cognitive difficulty in different representation forms, especially in the field of interaction design.Thus, Scientists pay attention to the design effectiveness based on visual perception. This study focuses on two problems: 1) The relationship between textual comprehension, spatial understanding and cognitive accuracy of text information; 2) The transformation differences of cognitive elements from text information to 3D image information. First, we conduct an experiment to show the cognitive transformation difference of text elements and 3D image elements. Then, we take the design of "Logoup" 3D modeling software (This is programming driven 3D modeling software) as an example, and applies the experimental results in this study to the interface design of the software. By setting up horizontal and vertical reference planes in the real-time rendering area of the software, we can improve the cognitive efficiency and user experience of users and provide non-professional 3D modeling skill of users with an entrance to create 3D shapes.

Keywords: Design cognition, Design methods, Design for interfaces, Spatial Understanding and Representation

Contact:
Li, Xuenan
East China University of Science and Technology, China
School of Art Design and Media
China, People's Republic of
xuenan332@sina.com

1 INTRODUCTION

Information is the crucial element of cognitive science (Bogdan, 1988). It reflects the relationship between objects, and is essentially the form of interaction between systems. Information often has different organizational patterns, because the world has different levels of composition under different constraints. Information can be divided into two forms from the perspective of cognition: material and semantic. The material information is formed by the arrangement of objects, attributes and time that are often organized organically due to their causality. On the other hand, the semantic information comes from people's more intuitive conceptual understanding of objects. Semantic information is a kind of material information with specific function and purpose, which is the result of human cognitive activities.

In recent years, with the development of computer graphics and human-computer interaction technology, dynamic interaction design has become commonplace. We can watch three-dimensional weather simulation on TV, or use Google Earth to pan and zoom any location of interest. Scientists in the field of cognition have always believed that the same information has different cognitive difficulty in different representation forms (Larkin, 1987, Breslow, 2009, Hegarty, 2010), which is more obvious in the field of interaction design. People pay attention to the design effectiveness based on visual perception in many different disciplines (Phadnis, 2019, Horvat, 2019). In the field of cartography, scholars pay attention to the problem of map design (Dent, 1985). Statisticians focus on how to better design charts to gain insight into data patterns (Cleveland, 1985, Tufte, 2001). Researchers in human engineering psychology are committed to developing design principles for process control, aviation, medicine and route planning (Smallman, 2005, Vicente, 2002). Researchers in the field of education focus on the development of dynamic interactive visual education platform to teach students different meanings of the same phenomenon (Ainsworth, 2006, Wu, 2001). Researchers in the fields of science visualization (McCormick, 1987), information visualization (Card, 1999) and geographic visualization (MacEachren, 2001) have focused on how to make the best use of new information technology to reveal complex data patterns. Sufficient design can help users avoid challenges on a cognitive level, guide them to interact with products smoothly. (Ning, 2019)

Two problems are discussed in this study: 1) The relationship between language comprehension, spatial understanding and cognitive accuracy of text information; 2) The transformation differences of cognitive elements between text information and image information. In this paper, we first conduct an experiment to show the cognitive transformation difference of text elements and 3D image elements. Then, the experimental results are used to guide the design of "Logoup" procedural modeling software, and improve the cognitive efficiency and interactive experience of users.

2 METHODOLOGY

2.1 A cognitive theory of multimedia learning

According to Mayer's cognitive theory (Mayer, 2002), the user possesses a visual information processing system and a verbal information processing, such that auditory narration goes into the verbal system whereas animation goes into the visual system.
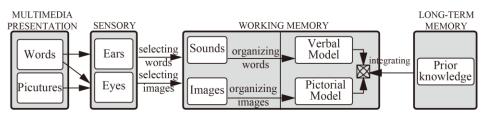

Figure 1. Mayer's cognitive theory of multimedia learning (Mayer, 1997)

Figure 1 presents Mayer's cognitive model of multimedia learning intended to represent the human information-processing system. The boxes represent memory stores, including sensory memory, working memory, and long-term memory. Sensory memory allows for pictures and words to be held as visual images for a short time period in eyes and for spoken words and other sounds to be held as auditory signals for a short time period in ears. Working memory is used for temporally holding and manipulating knowledge in active consciousness. Long-term memory can hold large amounts of

knowledge over long time periods, but to actively think about material in long-term memory it must be brought into working memory.

The information perceived by users can be summed up as words and pictures. Words refer to the text information based on certain grammars, which also includes some specific general symbols, such as symbols in mathematical, physical and chemical formulas. To exact text information, we need to understand its meaning through grammars, such as nouns, verbs and adjectives. For example, the first line for a case of illness is usually the patient's name, gender, age and other information; the second line is maybe the name and dosage of the prescribed drug; the last line is always the doctor's signature and signature date. The textual structure itself provides a lot of text information. Pictures refer to the image information composed of pixel units. The organization of image information is usually free, diverse and unstructured. To exact image information, users should map the pixel units in the image to existing knowledge stored in their brain, which may be the text that they need to explain.

Compared with image information, text information is a more direct way of information expression. The logic of text information is usually item oriented and meaningful, while the organization of image information is usually free and diversified. Therefore, image information is more difficult to extract than text information, but the expression of image information is more flexible than text information.
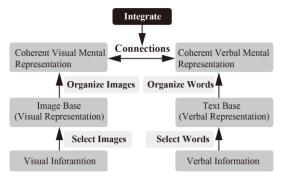

Figure 2. Generative theory of multimedia learning

The most crucial step in multimedia learning involves making connections between text and image information, as shown in Figure 2. This step involves two separate representations — a pictorial model to have an integrated visual representation and a verbal model to have an integrated verbal representation, in which the relations between text and image information are mapped from one model onto the other. In addition, this integrated step also include connections with prior knowledge.

2.2 Cognitive load experiment in multimedia learning

Multimedia learning involves the processing of text and image information, the corresponding verbal and spatial abilities are especially of interest. Plass (Plass, 2003) conducted a Vocabulary Test, a paper-and-pencil test in which the learner must determine the meaning of words printed in the left column on the page, to measure the speeded performance in processing a meaningful text, which is consistent with the verbal ability in multimedia learning. He also conducted a Card Rotation Test, a paper-and-pencil test in which the learner must determine whether two figures are the same or different, to measure the speeded performance in processing a meaningful image, which is consistent with the spatial ability in multimedia learning.

As shown in Figure 3, Plass's Vocabulary Test showed that: 1) When text annotation and image annotation exist at the same time, learners with high reading comprehension get the highest lexical accuracy. 2) When only text annotation was used, learners with low reading comprehension showed highest lexical accuracy. 3) Learners with high reading comprehension get generally higher lexical accuracy than learners with low reading comprehension. When text annotation and image annotation exist at the same time, or only image annotation exists, learners with low reading comprehension may be misled by image annotation.
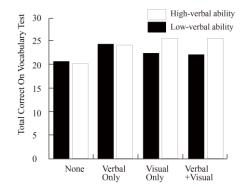
Figure 3. Total of correct answers on vocabulary test depending on treatment type for high-verbal ability learners and low-verbal ability learners

As shown in Figure 4, Plass 's Card Rotation Test showed that: 1) When only image annotation was used, learners with high spatial comprehension have the highest cognitive accuracy. 2) When text annotation and image annotation exist at the same time, learners with low spatial comprehension have the highest cognitive accuracy. When only image annotation is used, learners with low spatial comprehension may lack enough long-term memory to establish the relationship between image annotation and vocabulary. Therefore, in cognitive activities, learners with low spatial comprehension tend to use text information for cognition.
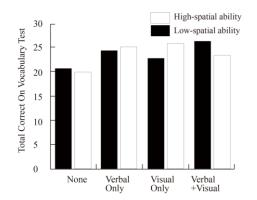

Figure 4. Total of correct answers on vocabulary test depending on treatment type for high-spatial ability learners and low-spatial ability learners.

3 COGNITIVE EXPERIMENT: TRANSFORMATION FROM TEXT TO 3D IMAGE INFORMATION

3.1 Verbal and spatial abilities

The cognitive efficiency of text and image information depends on people's verbal and spatial abilities. Verbal ability refers to the reading comprehension of users to extract text information from a meaningful text. Spatial ability refers to the spatial comprehension of users to extract image information from a meaningful image. Verbal ability can be measured by the so-called NDRT (the Nelson Denny reading test) method (Cummins, 1981), while spatial ability can be measured by the so-called CRT (card rotation test) method (Ekstrom, 1976).

Users' efficiency of cognitive transformation between text and 3D image information is one of the important factors to be considered in design activities. Plass's cognitive load experiment in multimedia learning as discussed in section 2.2 gave a convincing discussion on the relationship between verbal ability, spatial ability and cognitive accuracy of text information. However, he didn't discuss learners' efficiency of cognitive transformation between text and image information. In this paper, we conduct a cognitive difference experiment based on the transformation from text information to 3D image information to improve the design of 3D software interface.

3.2 Fast sketch experiment of cubes

Early studies in the field of visual cognition show that visual cognition in three-dimensional space can be decomposed into different visual components, including points, lines and surfaces distributed in the three-dimensional space (Uttal, 1995). Vision is perhaps the most important way for people to quickly recognize their surroundings. But how to measure people's visual cognition ability? According to Shoshi's research, the experimental method of fast sketch can reflect personalized cognitive styles and cognitive strategies, which is a way to express users' cognitive differences (Bar-Eli, 2013). The process of quick sketch can directly record the path of users' fleeting ideas (Tversky, 2002). Abidin said that the visual hand sketches point out certain meaning and can be categorized with respect to perceptual characteristics "(Abidin, 2011). There are profound differences between manual and automated morphing. Features and relations are identified which cumulatively generate a fuller understanding, or 'feel for' the configuration with which designer's is working (Abidin, 2011). Geometric elements are common materials used in cognitive experiments. Hermann used basic geometric shapes such as square, circle, triangle and crescent as the materials in his cognitive experiments, to study the sequence learning preference or cognitive habit of people in non-verbal state (Bulf, 2017). The use of linguistic interpretations is significant as a mean of analysis in order to examine formgiving (Abidin, 2008). Therefore, we adopt the experimental method based on fast sketch of cubes in following sections.

The experimental content should avoid human interference in the sketch drawing process, reduce the noise of experimental data, avoid human factors and environmental factors of sketch mode such as the participants' drawing ability, their professional background knowledge, the feasibility of the experimental content, and even the choice of drawing tools. In this experiment, participants were asked to quickly draw an arbitrary cube. Selecting "arbitrary cube" as the drawing content is mainly based on the following considerations:

1) The general acceptability of drawing content. First, the concept of cube is clear for participants, and it is easy for them to use single line segments to draw cubes. At the same time, the differences in the drawing process can be observed more clearly (the whole process of video analysis is used in the experiment). Second, the difficulty of drawing is more appropriate for participants, they can better focus on the experiment itself rather than whether the drawing is beautiful or not.

2) The word "cube" has no ambiguity when it is described as a spatial concept. It can completely represent the basic spatial information or spatial concept. At the same time, it has more abundant spatial information than circle, square and other two-dimensional graphics, and it is more operational than drawing a cartoon image, which eliminates the limitation of professional background knowledge. So this study is more related to the cognition of spatial information than drawing other shapes.

3.3 Experimental method

3.3.1 Hypotheses

In Plass's cognitive load experiment, learners only need to determine the meaning of words or images, which is a simple single-step decision process, so the accuracy of decision is the only evaluation metric of learners' cognitive ability. The fast sketch experiment of cubes in this study is far more complex. Learners need to transform the text information of the word "cube" to the image information of 3D cube images, by drawing the "cube" sketch using their hands and pens.

Intuitively, the subjects' gender, age, cognitive habits, and level of education all affect how well they draw the cube.Their gender and age may affect their personality and behavior, which in turn affects the time it takes them to draw the cube.Their cognitive habits and level of education may affect how they understand the concept of a cube, which in turn affects their preferred painting elevation and shape and size of the cubes they draw.

The process of fast sketching is a long-term and multi-step procedure, which can be described in the following aspects:

1) During the long-term and multi-step fast sketching process, learners can draw the points, lines and surfaces in different order, which reflects their different understanding of the concept of the word "cube".

2) The length of drawing time reflects the user's familiarity with the concept of the word "cube" and their painting ability, which is related to their educational background and prior knowledge.

3) The drawing results, including the size and shape of the cube, reflect their different understanding of the concept of the word "cube" and their painting ability.

All these three aspects are consistent with learners' efficiency of their cognitive transformation ability from text information (the word "cube") to image information (the hand painted 3D cube images). Therefore, we use these three aspects as evaluation metrics of learners' cognitive transformation ability in this study.

3.3.2 Participants

A total of 33 participants were randomly selected from different universities, including 42% undergraduates, 27% master students and 31% doctoral students. This includes 55% male and 45% female. Among them, 5 had received professional painting training, and 28 had no professional painting foundation. "Professional training" refers to the background of professional painting, art or spatial thinking training. Their age is about 20-30 years old, and they all have no relevant experimental experience.

3.3.3 Materials and procedure

A smart phone is mainly used to record the whole dynamic process from the beginning of drawing to the end of drawing. A black signature pen and an A4 copy paper are provided to draw sketches. In this experiment, each elevation of the cube is defined as Figure 5. XY plane is defined as X elevation, YZ plane is defined as Y elevation, ZX plane is defined as Z elevation.
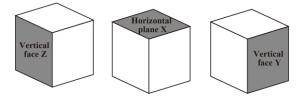

Figure 5. The definition of X, Y and Z elevations in the cube.

Each participant is asked to draw an arbitrary cube. The cube here refers to a regular polyhedron composed of six square faces, which is also called a regular hexahedron. It has 12 sides and 8 vertices. To avoid purposeful guidance or suggestion of test answers, any professional vocabulary or reference image is omitted in the questionnaire. This study focuses on the valuation metrics as described in section 3.3.1, and ignores other subjective factors in the drawing process.

3.4 Experimental results

3.4.1 Preferred painting elevation

During the process of drawing the "cube" sketch, learners with different understanding of the concept of the word "cube" will select the elevation that they prefer to draw first. In this experiment, 33 valid samples are included. The solid dots represent the cube elevation drawn by the participants first. As shown in Table 1, 57.6% of the participants preferred to draw X elevation. 42.4% of the participants chose to draw Z elevation first. No one chose to draw the Y elevation first. This may be related to the human preference or innate habit of sequential learning from left to right and from top to bottom, and it may also be related to the individual's reading habits, understanding habits, or physiological structure. In the experimental data, 42.4% of the participants preferred the left elevation as the reference plane to express the 3D space; 57.6% of the participants preferred the horizontal elevation as the reference plane to express the 3D space.

The results of this experiment have a key design significance for the research on the cognitive efficiency of the transformation from text information and 3D image information. In the process of transformation, there are differences in participants' spatial understanding and expression. X elevation (upper horizontal plane) and Z elevation (left plane) of cubes have obvious so-called first selection advantage. This advantage of "spatial orientation" can be used to improve the cognitive efficiency and experience of users.

Table 1. The difference of expression order of 3D cubes

Plane	Amount	Percentage (%)
X-Plane	19	57.6
Y-Plane	0	0
Z-Plane	14	42.4

3.4.2 The length of drawing time

In this experiment, the time spent by the participants to complete the drawing task was statistically analyzed. The individual drawing time represents the time of individual cognition in the process of transforming text information into 3D image information. Figure 6 shows the drawing time curves of participants with different educational levels, Figure 7 is the mean drawing time of participants with different educational levels. The average drawing time of undergraduate, master and doctor participants were 9.5 seconds, 9.3 seconds and 8.3 seconds, respectively (Figure 7). The t-value test of drawing time for participants with different educational background is shown in Table 2. There was no significant difference in drawing time among participants with different educational background ($P > 0.05$). The task of "Cube" drawing in this experiment had nothing to do with educational background, but only related to individual spatial comprehension and text comprehension. Figure 8 shows some participants' drawing processes which were recorded as video. The video can be clearly analyzed frame by frame.

In the experiment, the participants' drawing time is mainly distributed between 3 seconds and 20 seconds, showing obvious cognitive differences of text transformed images, which can be called individual differences in understanding and drawing "Cube".
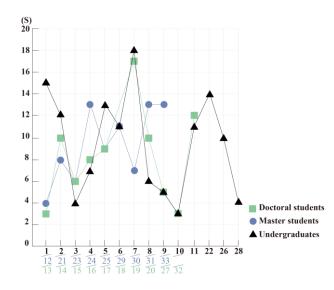

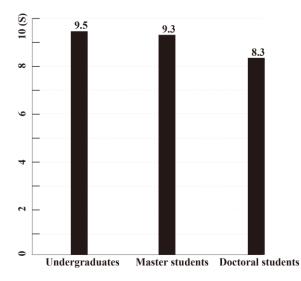Figure 6. Drawing time curves of participants with different educational levels.

Figure 7. Mean drawing time of participants with different educational levels.

Table 2. T-value test of Figure 6

	Undergraduate/Master	Master/Doctor	Undergraduate/Doctor
T-value	0.1144	0.5844	0.6644

Figure 8. Excerpts of participants' drawing process.

3.4.3 The size and shape of drawing cubes

In this experiment, participants were not constrained by the size of cubes, because the size constraint may potentially increase the difficulty of drawing. A piece of A4 paper (210 * 297mm) was provided for participants, which was a recessive constraint on the size.

During the drawing process, some participants drew the auxiliary structural dotted lines of the cube, while others did not. Correlation analysis of the relationship between drawing time and drawing size was divided into two groups (Figure 9 and Figure 10). There was no correlation between drawing time and drawing size (0.01) when structural dotted line was not drawn (Figure 9), while there was a positive correlation between drawing time and drawing size (0.28) when structural dotted line was drawn. Although there were significant differences in the drawing time of individuals, all participants showed consistent and unified drawing results.
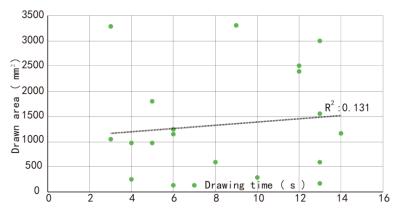

Figure 9. The relationship between size and duration when the dotted line is not drawn.
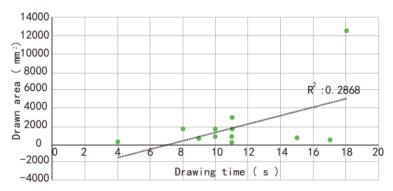
Figure 10. The relationship between size and duration when the dotted line is drawn.

4 DISCUSSION AND CONCLUSION

The basic theme of this study was that the transformation from text information to image information, which is the crucial content of cognitive learning, can potentially have deleterious effects depending on the individual differences of the learners. The experimental results in this study show that the horizontal and vertical reference planes should be established as far as possible when the 3D software interface involves spatial information, so the advantage of "spatial orientation" can be used to improve the cognitive efficiency and interactive experience of users.

Based on the experimental results in Section 3.4, we can draw the following conclusions.

(1) The experimental results demonstrate the cognitive difference when users need to transform the text information to image information. This cognitive difference is not only related to users' verbal ability and spatial ability which have been defined in multimedia learning theory and demonstrated in Plass's cognitive load experiment, but also related to users' behaviour habits, prior knowledge, education level and other factors as shown in this study.

(2) This study establishes the connection between user attributes (such as behaviour habits, prior knowledge, education level) and their cognitive transformation ability. First, 57.6% of the participants preferred to draw X elevation first. 42.4% of the participants chose to draw Z elevation first. No one chose to draw the Y elevation first. This may be related to the human preference or innate habit of sequential learning from left to right and from top to bottom. Second, there was no significant difference in drawing time among participants with different educational background, but the participants' drawing time is mainly distributed between 3 seconds and 20 seconds, showing obvious cognitive differences of text transformed images. Third, during the drawing process, some participants drew the auxiliary structural dotted lines of the cube, while others did not, there was a positive correlation between drawing time and drawing size (0.28) when structural dotted line was drawn.

(3) In the follow-up work of this study, we take the design of "Logoup" 3D modeling software (This is programming driven 3D modeling software) as an example, and applies the experimental results in this study to the interface design of the software. By setting up horizontal and vertical reference planes in the real-time rendering area of the software, we can improve the cognitive efficiency and user experience of users and provide non-professional 3D modeling skill of users with an entrance to create 3D shapes.

ACKNOWLEDGMENTS

REFERENCES

Bogdan R. Information and semantic cognition: An ontological account. Mind and Language, 1988, 3(2): 81-122. https://doi.org/10.1111/j.1468-0017.1988.tb00136.x

Larkin J H, Simon H A. Why a diagram is (sometimes) worth ten thousand words. Cognitive science, 1987, 11(1): 65-100. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

Breslow L A, Trafton J G, Ratwani R M. A perceptual process approach to selecting color scales for complex visualizations. Journal of Experimental Psychology: Applied, 2009, 15(1): 25. DOI: 10.1037/a0015085

Hegarty M, Canham M S, Fabrikant S I. Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 2010, 36(1): 37. https://dx.doi.org/10.1037/a0017683

Phadnis, V. S. , Leonardo, K. A. , Wallace, D. R. , & Olechowski, A. L. . (2019). An exploratory study comparing cad tools and working styles for implementing design changes. Proceedings of the Design Society International Conference on Engineering Design, 1(1), 1383-1392. DOI: 10.1017/dsi.2019.144

Horvat, N. , Kec, S. , Martinec, T. , Lukaevi, F. , & Perii, M. M. . (2019). Comparing virtual reality and desktop interface for reviewing 3d cad models. Proceedings of the Design Society International Conference on Engineering Design, 1(1), 1923-1932. DOI: 10.1017/dsi.2019.198

Dent B D, Torguson J S, Hodler T W. Cartography: Thematic map design. Boston: WCB/McGraw-Hill, 1999.

Cleveland W S, Cleveland W S. The elements of graphing data. Monterey, CA: Wadsworth Advanced Books and Software, 1985. https://doi.org/10.1016/0004-6981(86)90109-5

Tufte E R. The visual display of quantitative information. Cheshire, CT: Graphics press, 2001.

Smallman H S, John M S. Naïve realism: Misplaced faith in realistic displays. Ergonomics in design, 2005, 13(3): 6-13.

Vicente K J. Ecological interface design: Progress and challenges. Human factors, 2002, 44(1): 62-78. https://doi.org/10.1518/0018720024494829

Ainsworth S. DeFT: A conceptual framework for considering learning with multiple representations. Learning and instruction, 2006, 16(3): 183-198.

Wu H K, Krajcik J S, Soloway E. Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 2001, 38(7): 821-842. https://doi.org/10.1002/tea.1033

McCormick B H, DeFanti T, & Brown M D. Visualization and scientific computing. Arlington, VA:National Science Foundation, 1987.

Card M. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999. ISBN: 9781558605336

MacEachren A M, Kraak M J. Research challenges in geovisualization. Cartography and geographic information science, 2001, 28(1): 3-12. https://doi.org/10.1559/152304001782173970

Ning, W. , Goodman-Deane, J. , & Clarkson, P. J. . (2019). Addressing cognitive challenges in design – a review on existing approaches. Proceedings of the Design Society International Conference on Engineering Design, 1(1), 2775-2784. DOI: 10.1017/dsi.2019.284

Mayer R E. Multimedia learning[M]//Psychology of learning and motivation. Academic Press, 2002, 41: 37-38. http://dx.doi.org/10.1017/cbo9781139164603

Plass J L, Chun D M, Mayer R E, et al. Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities.Computers in Human Behavior, 2003, 19(2):221-243. http://dx.doi.org /10.1016/S0747-5632(02)00015-8

Cummins R P. Test review: The Nelson-Denny Reading Test (Forms E and F)[J]. Journal of Reading, 1981, 25(1): 54-59.

Ekstrom, R.B., French, J.W., Harman, H.H., & Dermen, D. Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service, 1976, 149-152.

Uttal W R, Baruch T, Allen L. The effect of combinations of image degradations in a discrimination task. Perception & Psychophysics, 1995, 57(5):668-681. DOI: 10.3758/BF03213272

Bar-Eli, Shoshi. Sketching profiles: Awareness to individual differences in sketching as a means of enhancing design solution development. Design Studies, 2013, 34(4):472-493. DOI: 10.1016/j.destud.2013.01.007

Tversky, B. What do sketches say about thinking. In Proc of Conf on AAAI Technical report, Alberta, 2002.

Abidin, S.Z., Warell, A., & Liem, A. (2011). The significance of form elements: A study of representational content of design sketches. In Proceedings of the DESIRE'11 Conference on Creativity and Innovation in Design (pp. 21-30). https://dx.doi.org/10.1145/2079216.2079219 .

Abidin, S.Z., Warell, A., & Liem, A. (2011). Understanding styling activity of automotive designers: A study of manual interpolative morphing through freehand sketching. In ICED 11 - 18th International Conference on Engineering Design - Impacting Society Through Engineering Design (PART 1 ed., Vol. 9, pp. 357-366).

Bulf H , De Hevia M D , Gariboldi V , et al. Infants learn better from left to right: a directional bias in infants' sequence learning[J]. Scientific Reports, 2017, 7(1):2437. DOI: 10.1038/s41598-017-02466-w.

Zainal Abidin, S., Sigurjónsson, J., Liem, A., & Keitsch, M. (2008). On the role of form giving in design. In DS 46: Proceedings of E and PDE 2008, the 10th International Conference on Engineering and Product Design Education.