Artificial Intelligence and Human-Computer Interaction
Y. Ye and H. Zhou (Eds.)

doi:10.3233/FAIA250172

Comparative Study of User Design Ideas in Generative Artificial Intelligence: An SEM Model Based on Multimodal Discourse Theory

Zhengtang TAN[a,b], Huixiang ZHANG[b], Zhuyun YUAN[b] and Fangzhou GU[c,1]
[a] *National Teaching Demonstration Center Laboratory, College of Engineering and Design, Hunan Normal University, Changsha, China*
[b] *College of Engineering and Design, Hunan Normal University, Changsha, China*
[c] *School of Art and Design, Changsha University, Changsha, China*
ORCiD ID: Zhengtang TAN https://orcid.org/0000-0002-0862-311X

Abstract. With the widespread application of generative artificial intelligence (AIGC) in design, its impact on design conception across different user groups is a significant research issue in design education. This study explores the differences in design conception between designer users and ordinary users using AIGC for poster design, and evaluates the application of multimodal discourse theory and SEM models. The study defines four variables in the poster design concept: representational meaning, interactive meaning, compositional meaning, and color meaning. A questionnaire with 22 items was distributed to 120 respondents. SEM analysis showed a good fit and significant inter-group differences. Designers' design ideation is influenced by all dimensions, while compositional meaning does not affect ordinary users. Designers' path coefficients were balanced across dimensions, while ordinary users showed significant differences, particularly in representational and interactive meaning. This study verifies the applicability of multimodal discourse theory and offers insights into design conception differences in the AIGC context. It emphasizes the need to improve non-professional users' design skills and provides guidance for developing more user-friendly AIGC tools.

Keywords. Multimodal discourse analysis; design ideation; generative artificial intelligence

1. Introduction

With the rapid development of artificial intelligence technology, generative artificial intelligence (AIGC) has demonstrated its innovative potential in many fields, especially in design and art creation [1]. AIGC tools, such as Midjourney, Disco Diffusion, and DALL-E 2, have not only promoted the diversification of creative expression but also accelerated the automation of the design process. The popularity of AIGC has greatly lowered the entry threshold of the design industry [2], allowing even users without professional design backgrounds to complete design tasks with the help of these tools.

[1] Corresponding Author: Fangzhou GU School of Art and Design, Changsha University, Changsha, China. E-mail: 512159229@qq.com.

The democratization of this technology not only enables non-professional designers to easily realize their creative ideas but also changes the employment structure and creative process of the design industry, making design more popular and personalized.

Design ideas have always been regarded as the core competitiveness of designers [3]. However, with the increasing popularity of AIGC technology, the difference in design expression between professional designers and ordinary users may be narrowing. By analyzing the task of poster design using AIGC tools, this study aims to explore whether professional designers' design ideation still shows obvious advantages in the current technological environment and how these differences affect the display of design quality and creativity.

In this article, we propose a structural equation model (SEM) based on multimodal discourse theory to measure the differences in design ideation among four variables: representational meaning, interactive meaning, compositional meaning, and color meaning in generative artificial intelligence tools. To explore the necessity of learning professional design knowledge, this paper first subdivided the thinking of poster design by introducing multimodal discourse theory and summarized 22 potential variables. Then, 22 related questions were designed for subsequent analysis by combining pictures and texts. After analyzing the 240 valid questionnaires obtained, we found significant differences in design conception between ordinary users and designer users.

2. Background

2.1. Design ideation

Design ideation is usually understood as design ideas and design intentions. Design ideation is both a method for innovatively solving design-based problems and a way to develop personal innovation capabilities. Design activities are an integral part of technical education as they are considered a key factor in solving the process of technological development [4]. Related research shows that design ideation plays a vital role in enabling designers to solve real-world problems in a creative and practical way [5]. Balakrishnan (2022) found that design ideation can help design students become more creative, enabling them to come up with innovative and useful designs. design ideation can also be an effective means of promoting creativity.

2.2. Multimodal discourse analysis theory

Multimodal discourse analysis theory refers to complex discourses that contain images, graphics, and animations in addition to text, or any text that uses more than one information code to convey information [6]. In his functional grammar, Halliday further proposed three meta-functions of language: the conceptual function, which expresses conceptual meaning; the interpersonal function, which expresses the relationship between the speaker and the listener, as well as the speaker's attitude towards the content; and the textual function, which conveys textual meaning.

Zhanzi L refined these three meta-functions and explained how to use Kress & Leeuwen's social semiotic method to analyze images, specifically in terms of representational meaning, interactive meaning, and compositional meaning [7]. Representational meaning refers to the relationship between the people, things, and places depicted in the image, encompassing both narrative and conceptual

representations. Interactive meaning mainly expresses the relationship between the elements that the graphic designer intends to convey—i.e., the subjects represented by the image—and the viewers of the image, also implying the potential attitudes of the information receivers towards these subjects. Compositional meaning concerns the overall arrangement of the image, relating how visual and interactive elements combine to form a coherent whole.

However, there are few studies showing how this method can directly assist in design, such as helping designers consider the weight ratio of different information codes during the design process.

2.3. Application of Structural Equation Modeling (SEM)in design

Structural equation modeling (SEM) is a statistical analysis tool that examines the relationships between variables based on their covariance matrix. It is widely used in design research to analyze user cognition and needs. Hernan Casakin and Shalom Levy highlighted that, although existing studies acknowledge the importance of design ability and creative thinking for professional design skills, a theoretical framework to explore the relationship between these elements in depth is still lacking [8].

Kuo-Liang Huang applied SEM in the context of artificial intelligence-generated content (AIGC) to investigate the potential role of various indicators in the design conception process. The advantage of SEM lies in its ability to directly observe measurement variables and manage multiple dependent variables [9]. Compared with traditional statistical analysis methods, SEM provides more reliable and precise results, making it a valuable tool for research in the design field.

2.4. Hypothesis

Based on the theory of multimodal discourse analysis, this study divides design ideation into four dimensions: representational meaning, interactive meaning, compositional meaning, and color meaning. The specific research methods will be elaborated in detail in the third part, "Methodology." On this basis, this study proposes the following two specific research questions:

- Q1: In the context of AIGC, are there significant differences in the design ideation of designer users and ordinary users in the four dimensions of representational meaning, interactive meaning, compositional meaning, and color meaning? This question will be verified through structural equation modeling, and the fit of the model and the significance of each path will be analyzed.
- Q2: In the context of AIGC, is there a significant difference in the importance of design ideation between designers and ordinary users in the design dimensions of representational meaning, interactive meaning, compositional meaning, and color meaning? We will use structural equation modeling (SEM) to analyze the numerical size and ordering of factor loadings of specific observed variables to evaluate the differences in the importance of design ideation in various design dimensions among different user groups.

3. Materials and methods

3.1. Methods

This study synthesizes multimodal discourse analysis theory and conducts an in-depth analysis of its application in the field of poster design. First, we used the group discussion method to screen research related to design ideation, and finally selected elements from multimodal discourse analysis theory including aesthetic value, information value, social value, process of action, process of reaction, process of analysis, symbolism, projection, eye contact, social distance, perspective of viewing, modal scene, information value, salience, and framing.

Based on the theory of visual grammar [10], this study further divides the design elements into three main analytical dimensions: representational meaning, interactive meaning, and compositional meaning. Representational meaning represents the relationship between the people, things, and places described in the image. Interactive meaning analyzes the relationship between the information conveyed by the designer through graphics and the viewer, emphasizing the audience's possible attitude towards the things represented by the image. Compositional meaning focuses on the overall structure of the image, exploring how visual and interactive elements interact to form a unified and meaningful whole [11].

To systematically refine the design elements, this study adopted the card sorting method. This method was jointly implemented by three doctoral students in the field of design and two master's students. The collected data was comprehensively sorted and summarized, resulting in the selection of 22 key design elements.

3.2. Materials

In the theory of multimodal discourse analysis, there are many professional terms that are not easy for users to understand. For example, the action process refers to the behavior of the participant sending a vector. To facilitate users' understanding of the subdivision of design ideation, we conducted discussions through an expert group consisting of 3 PhDs and 2 master students in the field of design. Using the picture-text completion method, 22 related questions were designed through picture book images and text. The questions are shown in Table1, and the final picture book image is shown in Figure 1. In this study, a 5-level Likert scale was used to score the degree of consideration of design factors in the design sample (1-not considered at all, 5-completely considered).

3.3. Process

Our data collection and analysis process involves three steps, as outlined below. First, we gathered questionnaires from target users using Wenjuanxing, a professional online survey tool. The survey primarily focused on students and graduates from Hunan University and Hunan Normal University, resulting in a total of 562 questionnaires. Of these, 301 were valid, comprising 176 from ordinary users and 145 from designer users. To ensure consistency between the two groups, we eventually sampled 120 responses from each group, with the sampling based on proportional distribution by age and educational level. The target users completed a demographic questionnaire (covering age, education, and gender) along with questions related to poster design (such as viewpoint, composition, and color tone) [12].

Table 1. Latent variables and observed variables.

Variable Category	Latent variables	Observed variables
Exogenous variables	Representational meaning	Eye contact is direct and clear (RM1)
		Reflect the phenomenon truthfully (RM2)
		Poster design is related to the expression of meaning (RM3)
		Add text description (RM4)
		The poster elements design must be within the scope of the poster theme (RM5)
	Interactive meaning	Establish direct eye contact between the poster subject and the viewer (IM1)
		Whether to use a top-down, top-down, or level-of-view perspective (IM2)
		Color details and authenticity fully restore reality (IM3)
		The poster's narrative scene completely restores reality (IM4)
		The main subject of the poster uses a close-up, medium shot or long shot (IM5)
	Compositional meaning	Whether to compile known information and new information (CM3)
		Whether to arrange core information and auxiliary information (CM4)
		Whether to arrange ideal information and actual information (CM7)
		Will the poster information be cut? (CM8)
		Whether to arrange foreground and background (CM9)
	Color meaning	Color tone tendency (CMO1)
		Saturation (CMO2)
		Brightness (CMO3)
		Contrast (CMO4)
Endogenous variables	Design ideation	Aesthetic value (DT1)
		Information value (DT2)
		Social Value (DT3)

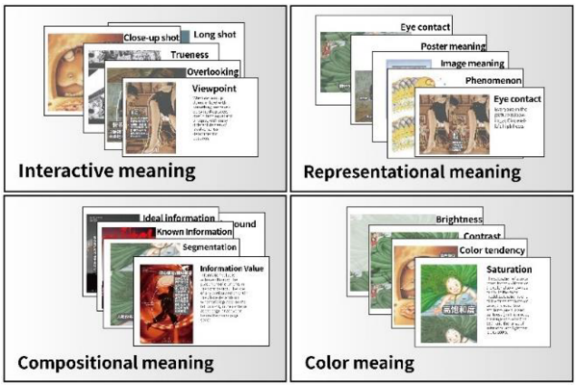

Figure 1. Picture question design based on multimodal discourse analysis.

Next, we conducted an empirical analysis of the theoretical model on the collected questionnaires. This was followed by descriptive statistical analysis, reliability and validity analysis. Finally, we performed descriptive statistics and normality tests to ensure the credibility of the model results. Lastly, we constructed and compared the structural equation models (SEM) of the two types of users.

4. Analyze

4.1. Descriptive statistics

We used SPSS 26.0 software to conduct descriptive statistical analysis on the data. The analysis results show that the proportion of ordinary users and designers among the participants is close, 49.5% and 50.4%, respectively. Among the ordinary user group, the age distribution is mainly concentrated in the 19-25 years old range, accounting for

85.8%, and the education level is mainly college undergraduate, accounting for as high as 92.5%. The age distribution of the designer group is also mainly in the 19-25 years old range, accounting for 76.2%, with the educational background primarily being college undergraduate, accounting for 73.0%.

4.2. Reliability analysis

The reliability coefficients for design ideation, representational meaning, interactive meaning, compositional meaning, and color meaning for both designer users and ordinary users are as follows: 0.861, 0.902, 0.927, 0.899, 0.899 for designer users, and 0.893, 0.884, 0.934, 0.923, 0.897 for ordinary users. The reliability coefficients of multimodal discourse analysis, design ideation, and each secondary dimension all fall within the range of 0.6 to 1.0. Therefore, these results indicate that the scales used in this study have good internal consistency.

4.3. Related analysis

In this analysis, Pearson correlation analysis was used to conduct an exploratory analysis of the correlation between various variables. According to the analysis results, it can be seen that there is a positive correlation between each variable, and they are all significant at the 99% significance level [13].

5. Results

After running the model, we obtained the standardized coefficients for both designer users and ordinary users, as shown in Table 2. The final SEM model structure is illustrated in Figure 2. In Table 2, the paths DT, RM, IM, CM, and COM refer to design ideation, representational meaning, interactive meaning, compositional meaning, and color meaning, respectively. "DT<-RM," "DT<-IM," "DT<-CM," and "DT<-COM" represent the mutual influences among design ideation, representational meaning, interactive meaning, compositional meaning, and color meaning. For example, DT<-RM refers to the path coefficient of the influence of representational meaning on design ideation. We will describe the results for designer users and ordinary users separately.

5.1. Designer user results

From the analysis results in Table 2, it can be found that the standardized coefficients of all paths are greater than 0, and the corresponding P values are less than 0.05. The results show that the representational meaning, interactive meaning, compositional meaning, and color meaning in the multimodal discourse analysis theory have a significant positive impact on designers' design ideation [14]. There is a high correlation (0.615) between interactive meaning and representational meaning for designers.

Specifically, interactive meaning has the most significant impact on design ideation, with a standardized coefficient of 0.25. This is followed by representational meaning and compositional meaning, both with influence coefficients of 0.23, while the influence coefficient of color meaning is 0.20. Further analysis shows that IM2 has the greatest impact within interactive meaning, with a standardized coefficient of 0.93, followed by

IM3 with a standardized coefficient of 0.92. The impact coefficient of IM1 is 0.84, ranking third; the impact coefficients of IM4 and IM5 are 0.80 and 0.74, respectively, ranking fourth and fifth. Among the specific effects of representational meaning, RM2 has the largest effect, with a coefficient of 0.86, followed by RM1 with a coefficient of 0.85. RM5 ranks third with a coefficient of 0.818, while RM3 and RM4 rank fourth and fifth with coefficients of 0.778 and 0.736, respectively.

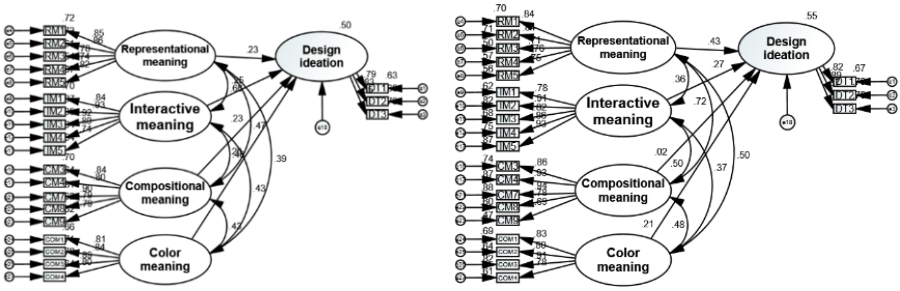

Figure 2. (A) SEM model results for designer users; (B) SEM model results for ordinary users.

Table 2. Path normalization coefficient.

User category	path	Standardized coefficient (β)	Standard Error	critical ratio	Significance(P)
Designer users	DT<-RM	0.230	0.123	1.965	0.049 (*)
	DT<-IM	0.252	0.124	2.136	0.033 (*)
	DT<-CM	0.228	0.089	2.298	0.022(*)
	DT<-COM	0.196	0.076	2.080	0.038(*)
Ordinary user	DT<-RM	0.429	0.150	3.300	***
	DT<-IM	0.274	0.070	3.057	0.002 (**)
	DT<-CM	0.023	0.102	0.184	0.857
	DT<-COM	0.212	0.085	2.272	0.023 (*)

(Note: * indicates 0.01< P <0.05; *** indicates P <0.001)

5.2. Ordinary user results

The normalized coefficients of each path for ordinary users are greater than 0. Specifically, the influence coefficient of representational meaning is 0.43, the influence coefficient of interactive meaning is 0.27, and the influence coefficient of color meaning is 0.21. The P values of these three are all less than 0.05, indicating that they have a significant positive impact on the design ideation of ordinary users. However, the standardized coefficient of compositional meaning is only 0.023, and the corresponding P value is 0.857, showing that its impact on the design ideation of ordinary users is not significant [15].

Looking at the data in Table 2 in further detail, within representational meaning, RM2 has the greatest impact, with a standardized coefficient of 0.843, followed by RM1 with a coefficient of 0.836. The coefficients of RM4, RM5, and RM3 are 0.757, 0.748, and 0.708, respectively. In terms of interactive meaning, IM5 has the most significant impact, with a coefficient of 0.935, followed by IM2 with a coefficient of 0.907. The coefficients of IM4, IM3, and IM1 are 0.864, 0.822, and 0.785, respectively. The correlation between representational meaning and compositional meaning is strongest among ordinary users (0.656).

6. Discussion

In the context of AIGC, design conception still plays a key role in design education to cultivate designers' design capabilities. Based on the results of multimodal discourse analysis theory, we found the following:
First, there are significant differences in the design ideation of designer users and ordinary users in the four dimensions of representational meaning, interactive meaning, compositional meaning, and color meaning. Second, there are significant differences in the importance of design ideation between designers and ordinary users in terms of design dimensions such as representational meaning, interactive meaning, compositional meaning, and color meaning.

6.1. Comparative discussion of user design ideation based on saliency

When comparing the significance data of designers and ordinary users in Table 2, we noticed that the path coefficients for designers in the four dimensions all have a significant impact on design ideation, while the compositional meaning for ordinary users does not show a significant impact ($P<0.001$). This indicates that designers can comprehensively consider design issues after receiving professional training in design creativity and ability, which is consistent with Cross's (2004) assertion that design ideation can improve designers' professional knowledge [16]. design ideation prompts designers to continuously accumulate professional knowledge to deal with unclear subjective design problems [17]. In contrast, the lack of prominence in compositional meaning among ordinary users may stem from a lack of professional skills training, which may lead to the neglect of effective layout of visual information in the design process.

6.2. Comparative discussion of user design ideation based on path coefficient

Designers' path coefficients in representational meaning, interactive meaning, compositional meaning, and color meaning show smaller differences, reflecting their balanced approach and professional training in the design process. This indicates their holistic thinking and problem-solving ability. In contrast, ordinary users show larger path coefficient differences in these dimensions, especially the significant effect of representational meaning, which may be related to their lesser design experience and technical knowledge. Interactive meaning showed the most significant impact among designer users, emphasizing their sensitivity in the layout of visual information to enhance interaction with the audience [18]. The results show that designers need to not only pay attention to the clarity of visual information but also ensure its understandability. This is crucial for designers when choosing aspects such as style, color, function, and texture to convey information [19].

This study focuses on the study of design ideation and also shows from the side that the weight ratio of design considerations differs between the two groups. We believe that ordinary users, when using AIGC for design and creation, should pay more attention to the layout of visual information to improve the quality of AI-generated works. Based on our combined picture and text questionnaire, visual aids can better help ordinary users understand layout principles [20]. On the other hand, the research found that although the weight ratio of the impact of design ideation differs between designers and ordinary users, and compositional meaning does not significantly affect ordinary users, the

essence of the principle of visual communication—encoding and reception of information—is reflected in both groups' consideration weights [21].

6.3. Limitations

This study employs a questionnaire survey combined with a structural equation model (SEM) to analyze the design concepts of designers and ordinary users. However, the questionnaire's reliance solely on self-reported data may introduce bias and may not accurately capture the participants' thought processes during design activities. This approach might overlook the subtle nuances of how users interact with AI tools during the design ideation process. Additionally, the study's focus on a specific task (poster design) limits its applicability, making it difficult to generalize the findings to other design tasks or contexts.

7. Conclusions

This paper introduced the theory of multimodal discourse analysis into the SEM framework to explore the differences in design ideation between designers and ordinary users, confirming the possibility of using SEM to interpret these differences [22]. In the specific implementation process, the questionnaire format was innovatively combined with pictures and texts to allow the subjects to understand the questions more intuitively and efficiently.

The results show that, compared to using qualitative analysis to study design ideation, SEM can draw more convincing conclusions through rigorous theoretical support and data analysis [23]. The conclusions drawn about design ideation between the two groups are as follows:

First. There is a small difference in the path coefficients of designers' influence on design ideation regarding representational meaning, interactive meaning, compositional meaning, and color meaning. This indicates the designer's multifaceted understanding of design needs, mainly reflected in the proficiency of professional skills, the innovation of design ideation, and the balance of design performance. Second. Representational meaning among ordinary users has a significant impact on the design ideation path coefficient. This suggests that ordinary users, even without professional training, will consider the completeness of information elements in the image, such as the relationship between people, things, and places. Third. Among ordinary users, compositional meaning has no significant impact on the design ideation path coefficient. Since ordinary users lack relevant professional knowledge and skills, the emphasis on the arrangement of design elements is not significant. This indicates the need for ordinary users to appropriately improve the arrangement and adjustment of design elements when using generative intelligence software.

Acknowledgment

We would like to thank the support of Hunan Province Social Science Foun- dation, Project: “Social Aesthetic Education Design of Huxiang Handicrafts Based on Children’s Embodied Cognition Perspective”, Project number: 21YBQ036.

References

[1] Yu H., Dong Y., Wu Q. User-centric AIGC products: Explainable Artificial Intelligence and AIGC products[J/OL]. (2023) 1-4. DOI: 10.48550/arXiv.2308.09877.

[2] Lu, Y., Guo, C., Dou, Y. et al. Could ChatGPT Imagine: Content Control for Artistic Painting Generation Via Large Language Models. J Intell Robot Syst (2023). 109, 39. DOI: 10.1007/s10846-023-01956-6.

[3] Pande M., Bharathi S. V. Theoretical foundations of design thinking – A constructivism learning approach to design thinking[J/OL]. Thinking Skills and Creativity, (2020). 36: 100637. DOI: 10.1016/j.tsc.2020.100637.

[4] Lewis, T. Design and inquiry: Bases for an accommodation between science and technology education in the curriculum? Journal of Research in Science Teaching, (2006). 43(3), 255–281. DOI: 10.1002/tea.20111.

[5] Balakrishnan, B. Exploring the impact of design thinking tool among design undergraduates: A study on creative skills and motivation to think creatively. International Journal of Technology and Design Education, (2022). 32(3), 1799–1812. DOI: 10.1007/s10798-021-09652-y.

[6] Halliday, M. A. K., Matthiessen, C. M. I. M. Halliday's introduction to functional grammar. Routledge. (2013). DOI: 10.4324/9780203431269

[7] Zhanzi, L. Social semiotic approach to multimodal discourse. Foreign Languages Research, (2003). 5: 1-8. Link.

[8] Hernan Casakin & Shalom Levy. Ideation and Design Ability as Antecedents for Design Expertise. Creativity Research Journal. (2020) DOI: 10.1080/10400419.2020.18347 42.

[9] Huang, K.-L., Liu, Y.-C., & Dong, M.-Q. Incorporating AIGC into design ideation: A study on self-efficacy and learning experience acceptance under higher-order thinking. Thinking Skills and Creativity, (2024). 52, 101508. DOI: 10.1016/j.tsc.2024.101508.

[10] Kress, G., & Leeuwen, T. V. Reading Images: The Grammar of Visual Design (3rd ed.). Routledge. (2020). DOI: 10.4324/9781003099857.

[11] Bezemer, J., & Kress, G. Writing in Multimodal Texts: A Social Semiotic Account of Designs for Learning. Writ. Commun. (2008). 25, 166–195. DOI: 10.1177/0741088307313177.

[12] Krosnick, J.A. Questionnaire Design. In: Vannette, D., Krosnick, J. (eds) The Palgrave Handbook of Survey Research. Palgrave Macmillan, Cham. (2018). DOI: 10.1007/978-3-319-54395-6_53.

[13] Cleophas, T.J., & Zwinderman, A.H. Bayesian Pearson Correlation Analysis. In: Modern Bayesian Statistics in Clinical Research. Springer, Cham. (2018). DOI: 10.1007/978-3-319-92747-3_11.

[14] Cheung, M. When mind, heart, and hands meet: communication design and designers. Int J Technol Des Educ, (2012) 22, 489–511. DOI: 10.1007/s10798-011-9155-7.

[15] Krosnick, J.A. (2018). Questionnaire Design. In: Vannette, D.L., & Krosnick, J.A. (Eds.), The Palgrave Handbook of Survey Research. Springer International Publishing, Cham, pp. 439–455. DOI: 10.1007/978-3-319-54395-6_53.

[16] Cross, N. G. The nature and nurture of design ability. Design Studies, (1990). 11(3), 127–140. DOI: 10.1016/0142-694X (90)90002-T.

[17] Cross, N. Design thinking: Understanding how designers think and work. London, UK: Berg Publishers. (2011). Link.

[18] Samara, T. Graphic Designer's Essential Reference: Visual Elements, Techniques, and Layout Strategies for Busy Designers. Rockport Publishers. (2011). Link.

[19] Boers, M. Designing effective graphs to get your message across. Annals of the Rheumatic Diseases, (2018). 77, 833-839. DOI: 10.1136/annrheumdis-2018-213356.

[20] Bajo, M. T. Semantic facilitation with pictures and words. Journal of Experimental Psychology: Learning, Memory, and Cognition, (1988). 14(4), 579. DOI: 10.1037/0278-7393.14.4.579.

[21] Biocca, F., Owen, C., Tang, A., & Bohil, C. Attention Issues in Spatial Information Systems: Directing Mobile Users' Visual Attention Using Augmented Reality. Journal of Management Information Systems, (2007). 23(4), 163–184. DOI: 10.2753/MIS0742-1222230408.

[22] Khong, P. W., & Song, J. P. Exploring user's emotional relationships with IT products: a structural equation model. In: Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces, DPPI '03. Association for Computing Machinery, New York, NY, USA, (2003). pp. 45–50. DOI: 10.1145/782896.782908.

[23] Bollen, K. A., & Noble, M. D. Structural equation models and the quantification of behavior. Proceedings of the National Academy of Sciences, (2011). 108(supplement_3), 15639-15646. DOI: 10.1073/pnas.1010661108.