Reconsidering Design Pedagogy through Diffusion Models

Selen Çiçek[1], Gözde Damla Turhan[2], Mine Özkar[3]
[1,3]Istanbul Technical University [2]Izmir University of Economics
[1,3]{cicekse20|ozkar}@itu.edu.tr [2]gozde.turhan@iue.edu.tr

The text-to-image based diffusion models are deep learning models that generate images from text-based narratives in user-generated prompts. These models use natural language processing (NLP) techniques to recognize narratives and generate corresponding images. This study associates the assignment-based learning-by-doing of design studio with the prompt-based diffusion models that require fine-tuning in each image generation. The reference is a specific formal education setup developed within the context of compulsory courses in design programs' curricula. We explore the implications of diffusion models for a model of the basic design studio as a case study. The term basic design implies a core and foundational element of design. To explore and evaluate the potential of AI tools to improve novice designers' design problem solving capabilities, a retrospective analysis was conducted for a series of basic design studio assignments. The first step of the study was to reframe the assignment briefs as design problems and student design works as design solutions. The outcomes of the identification were further used as input data to generate synthetic design solutions by text-to-image diffusion models. In the third step, the design solution sets generated by students and the diffusion models were comparatively assessed by design experts with regards to how well they answered to the design problems defined in the briefs. The initial findings showed that diffusion models were able to generate a myriad of design solutions in a short time. It is conjectured that this might help students to easily understand the ill-defined design problem requirements and generate visual concepts based on written descriptions. However, the comparison indicated the value of design reasoning conveyed in the studio, as it gets highlighted with the lack of improvement in the learning curve of the diffusion model recorded through the synthetic design process.

Keywords: *Deep Learning, Diffusion Models, Design Education, Basic Design, Design Problems*

INTRODUCTION

As the recent advancements in Deep Learning (DL) are shaping the design realm, it is inevitable to see its applications in design education (Schmidhuber, 2015). Although the integration of generative algorithms as pedagogical tools in design have been discussed widely, the potentials of text-to-image Diffusion Models (DM) is still an ongoing inquiry.

The inherent Natural Language Processing (NLP) techniques used in DM models recognize the narratives and generate corresponding images (Zbinden, 2022; Torfi et al., 2020; Nanda et al., 2020). In this study, the DM computer architecture is assumed suitable to assignment-based learning-by-doing type of design education as it incorporates solution assessment and development during a generative process. Therefore, within the scope of this paper, we aim to explore and discuss the potential of integrating DM models to the first-year of design education. The potential of text-to-image DMs in a first-year design studio is not only for image generation to enlarge the design solution space. It may also be used for reiterating and redefining the design spaces created by the students.

To explore DM models in a formal education setup, the basic design studio is selected as a case study. Basic design is a foundation studio and its selection as a case is significant due to its capacity in providing a ground for the students in their first encounters with the ambiguous nature of design problems. We hypothesize that a DM integration to a basic design studio could assist the student's comprehension of the implicit definitions of the problems, abstract concepts, and basic principles of design.

Our study is designed as a comparative assessment of two sets of solution spaces of design problems generated by the students and by a DM.

As discussed in the following sections in detail, a series of basic design studio assignments are analyzed retrospectively. We refer to the assignment briefs as design problems and to the student works as design solutions. We identify the ill- and well-defined problems of the design briefs. These identifications are used in creating the text-prompts to the DM. The solution spaces generated by students and DM are evaluated by semi-structured interviews with design experts.

First Encounters with Design Problems

The challenge for first year design students is initially with how design problems are defined (Ghom and George, 2020). The novice designer who has previously been exposed to well-defined or tamed problems during their former education, generally structured with one goal and one valid answer (Rittel and Webber, 1973; Coyne, 2005). Being used to instant solutions by pre-established methods and formulas in primary and secondary education, students are not accustomed to question the nature of the given problem or develop a methodology to tackle with it (Saranlı, 1998).

Since this first encounter with the ill-defined or wicked design problems generally has a devastating impact on novice students (Ustaomeroglu et al., 2015), the AI aided tools could help them to develop a design reasoning mechanism. The synthetic solution space generated by DM can allude to the see-move-see pattern for students (Schön, 1992a). It would enable them to deliver their organic design solutions, through their interpretations of the briefs.

Text-to-image Diffusion Models

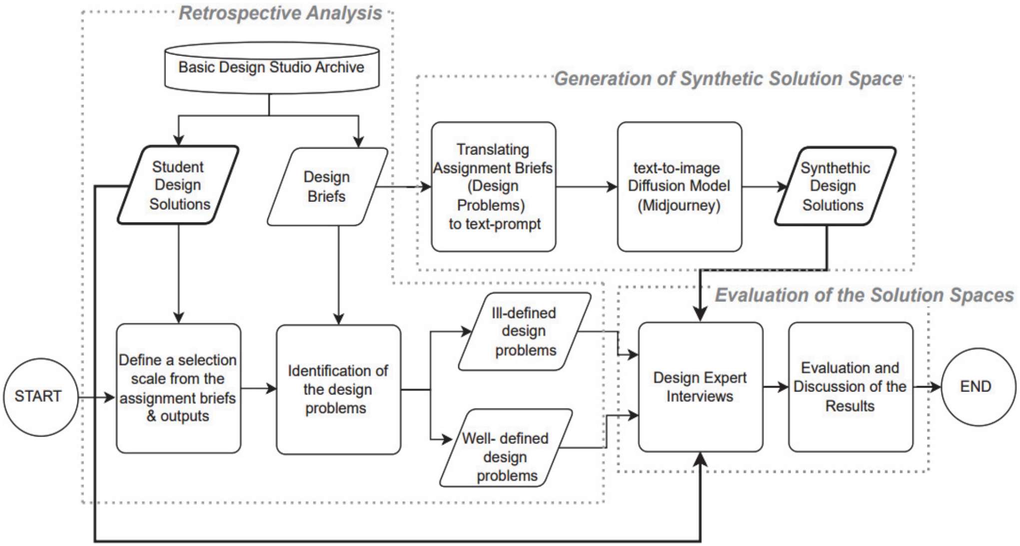
In the current state of the art, DMs stand out among other deep generative models for being capable of generating high-quality image samples in a multi-modal working environment (Dhariwal and Nichol, 2021). Implementing DM in a variety of different contexts and tasks is possible. As Vashani et al. (2017) states, the attention mechanism implemented in the transformer network of NLP model enables the text-to-image DM models to focus on various parts of the input text during encoding, allowing it to capture long-range dependencies between words in a sentence. Moreover, the self-attention mechanism used in the transformer architecture allows the model to compute attention weights for each word in the input sequence, which are used to weight the contributions of different words in computing the output representation (Vashani et al., 2017). Therefore, it is possible to generate images that closely match the input text description, even for the complex and nuanced descriptions.

The selection of the DM model for the scope of the study is due to the need to generate solutions by using the complex and nuanced descriptions of the briefs as text-prompts.

METHODOLOGY

In order to frame a formal setup for the case of the basic design studio, the suggested methodology

Figure 1
Study flowchart.



encompasses three main stages: a retrospective analysis of assignment briefs and student works to identify the inherent ill- and well-defined design problems; generating synthetic solutions by prompting design briefs into diffusion models; assessment of the student design works and synthetic solutions in design expert interviews (Figure 1).

Retrospective Analysis of a Basic Design Studio

To conduct the research, personal archives of a basic design studio from the Middle East Technical University ARCH 101 course is retrospectively analyzed. The archive includes the digital photographs of student works and the design briefs.

The student works in the archive were a selection by the design studio instructors and are exclusively examples deemed successful or of higher value than others in terms of the composition quality.

Four consecutive assignment briefs and the corresponding student solutions of the 2004-2005 fall semester are selected as the sample set to explore.

Each assignment is short and most of the time a two-stepped exercise. The brief for each describes the task and the quantitative limitations and material specification.

The initial analysis of the data indicated that in all semesters most of the assignments are given in a sequential order as multiple assignments and stages to define a continuous design process throughout the semester. The first criterion for the sampling was the availability of the photographs of student works corresponding to each assignment and its intermediate stages.

Identifying the Explicitness of the Design Problems

The series of assignments starts with a task of transforming certain sounds into a physical medium. The representations are then to be used as an element of the compositions in the following stages. Although, the problem starts to get concrete in terms of the physical medium, what is expected from the students is an interpretation of the abstract concepts to develop a design solution.

In our study, the problems in the briefs are identified as ill- or well-defined with regards to the literature (Simon, 2019; Cross, 1984; Rittel and Weber, 1973):

- The definitions that give information about the design elements material properties (i.e., material type, number, size, color), and representation mediums material properties are considered as well-defined problems.
- The statements that define the tasks with abstract concepts and notions that challenges the novice designer's comprehension level are considered as ill-defined. Since they do not define explicit directional steps, each designer can generate a myriad of solution way for the problem.

Table 1 displays the identification of the problems stated in the assignment briefs.

Generating Design Solutions by a Diffusion Model

In our study, the DM of the Midjourney research lab was used to generate solutions to the design problems given in the assignment briefs. It stands from other current state-of-the-art diffusion models *i.e.,* Dall-E 2, Stable Diffusion as being capable of enabling image prompting (Midjourney, 2023; OpenAI, 2023; Stability AI, 2023). Since the analyzed assignment briefs describe a sequential order of solution generation, the image instances of the former stages were given as the image prompts for the following stages generation procedures. The user interface is easy to use and publicly available from the Discord server. The design problems stated in the assignment briefs were translated into text-

A	S	Ill Defined Problems	Well- Defined Problems
2	1	number of sounds with a single syllabus / representing a sound on human body	selecting three sounds/positions / printed in black and white / on A4 sheets
	2	-	2D / half size photos/ 6 photos to be composed
3	1	transforming into measurable, reproducible geometric entities / regulating lines	enlarging two photographs / two different backgrounds / overlaying A3 size sketch paper on photographs
	2	translating on to a new medium / Obtain number of shapes by extracting	cartridge paper/ black for one and gray for the other
4	1	-	29.7x29.7cm cartridge paper (square canvas)
	2	regulating lines / act as one unit of the coming stage	21.2x21.2 cm white cartridge paper / nine copies
	3	the relation of the square composition to the whole / parts make a whole that is more than what they can achieve by themselves	nine copies of the square / 63.6x63.6 cm
5	3	contrast as the most perceptible quality	paper-tape of 12 mm / black A3-size cartridge paper / elements should start and end only at the edges

Table 1
Identified design problems. (A: Assignment No, S: Stage No.)

prompts by preserving the semantic organization of the description. It was crucial not to alter the semantic structure of written descriptions during the translation in order to make a comparative analysis between the design solution spaces produced by students and DM. However, since the DM requires to have a certain synthetic and semantic hierarchy within the text-prompts, the design briefs were altered in several ways as explained in the following section.

Translating the Design Problems into Text Prompts

The alterations made during the translation of the assignment briefs to text-prompts are grouped under three headings.

Syntactic Alterations. First group of alterations is based on the syntax and the sentence case of the written statements. In the Midjourney model, the text prompts get separated by commas in order to maintain the semantic hierarchy of the prompt (Midjourney, 2022). When the assignment tasks are analyzed, it became visible that comma usage was frequent. Therefore, to hold on to the semantic hierarchy of the brief definition, these commas are either eliminated, or replaced with “and”. If the brief is defined as a paragraph that contains two or three sentences, full stop sign is used in order not to interrupt description of the task process. Secondly, the DM prompting is with numerical descriptions by using verbal descriptions rather than numbers (Midjourney, 2023). So, in the cases where the design brief defines numerical specifications such as the number of design elements or the repetitions in the composition, these are transformed into verbal descriptions (Table 2).

Sentence Case Alterations. The expressions stated as briefing the students with the usage of personal pronouns or indications i.e. “you”, “your”, “a group of students” are eliminated. Because during the experimentation, it is noticed that the pronouns used in the prompts, deflects the model to depict the designer instead of depicting the design work itself. The Table 3 illustrates the mentioned deflection of the model by prompting with the pronouns.

Secondly, the experimentations on the model show that the material specifications defined at the end of text-prompt succeeds to generate instances that fit better into the context. These specifications were defined as the reference for the 2D design medium’s material aspects (Table 2).

Table 2
Syntactic and sentence-case alterations’ impact on the generated images visuals during translation process of briefs to text-prompts

	Original Brief	Text Prompt Translation
Text-prompts	Using paper-tape of twelve mm you are asked to make a composition in which contrast is the most perceptible quality. Your paper-tape elements should start and end only at the edges of the cartridge paper and they should not be adjacent.	Using twelve mm paper-tape make a composition in which contrast is the most perceptible quality. Paper-tape should start and end only at the edges of the cartridge paper, and they should not be adjacent, black cartridge paper, paper tape - -v 3
DM Solution		

The Use of Parameters and Versions. The size of the design medium can be defined by using specific parameter, aspect ratio (Midjourney, 2023). Since the aspect ratio defined by the default mode of image generation is 1:1, "--ar x: y" parameter is used at the end of the prompt cases in order to maintain the correct dimension of the design medium. The third version of the model (v3, entered the text prompt as "-- v3") used to generate all the instances. The parameters that control the image outputs visual characteristics, quality and style i.e stylize, chaos etc. kept as default in all generation cases.

Assessment of the Two Sets of Design Solutions

To evaluate the performance of two sets of the design solution spaces, respectively generated by students and the DM, a series of interviews were held with eight design experts.

All participants were either currently teaching in a first-year studio, or had experience in instructing in one. In order to avoid any biases that might occur for the design experts, all eight experts are selected from an institution different than METU.

All experts chosen for the study are currently teaching at Izmir University of Economics, Faculty of Fine Arts and Design. Since the basic design studio of the institution is situated as core studio for all design departments in the faculty, the professional backgrounds of the participants varied in terms of the design disciplines they belong into, including architects, interior architects, fashion designers, industrial designers, and visual communication designers.

The visual material from Solution Spaces A (SSA) and Solution Spaces B (SSB) were distributed to each participant, in addition to the assignment briefs. The SSA contains the original design works of students, whereas SBB contains the diffusion model generated solutions (Table 3).

However, how the SSA and SSB produced was not made explicit to the interviewees to avoid any bias might occur during the assessment. It was

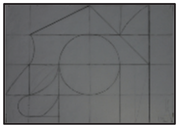
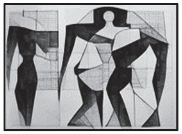
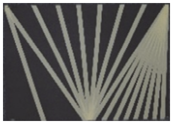
	S	SSA	SSB
2	1		
	2		
3	1		
	2		
4	1		
	2		
5	3		

Table 3
The student design solutions (SSA) and the synthetic solutions (SSB).

recorded that none of the participants anticipated the generation method of SSB.

Sessions were held separately in order to avoid reviewers influence each other. The participants were asked to rate each output in an ordinal scale (1 = poor; 2 = poor-average; 3 = average; 4 = average-excellent; 5 = excellent) for the four questions posed:

1. Please rate each design solution space's success to answer the given assignment brief in general.

2. Please rate each design solution space's success in terms of implying the well-defined problems on the design solution visually.

3. Please rate each design solution space's success in terms of implying the ill-defined problems on the design solution visually.

4. Please rate the overall performance of Group A and B, in terms of interconnectedness of the solutions through the process, defined by sequential assignment briefs and stages.

RESULTS AND DISCUSSION

For each question asked in the assessment process, the performance of the solutions is recorded separately for each reviewer.

Table 4 displays an overview of the reviewers' ratings for each solution space (SSA and SSB) under the four questions posed. To calculate each design solution space performance, the average of the points given by independent researchers in 1-5 scale converted into hundred percentage to increase the legibility of the results in Figure 2.

The overall performance results indicate that the student design solutions are slightly better in answering the requirements of the assignment briefs in general.

Considering the first questions' aim as evaluating of the performance of SSA and SSB in answering both ill- and well-defined problems, the result can be considered as an average of the second and third. However, the individual results of questions 2, 3 and 4 highlight a significant difference, that is aimed to be explored as a result of the study.

The Impact of Well-defined Problems

The second question aimed to evaluate the solution spaces' performance in answering the well-defined problems, in terms of the design medium and material specifications.

The result implies a noticeable difference of 21% between the solution spaces in terms of answering to the well-defined design problems stated in the briefs. As in the scope of the study these problems reframed as the material properties of the design elements and representation media. The apparent difference can be considered as an impact of a limitation of the study as, context-free training

Table 4
Assessment Results (Q: Questions, SSA: Solution Space A -student works, SSB: Solution Space B- DM works, R: Reviewers, M: Arithmetic mean

	Q 1		Q2		Q3		Q4	
R.	SSA	SSB	SSA	SSB	SSA	SSB	SSA	SSB
1	3,3	1,4	3,3	1,3	3,3	1,3	4	3
2	3,1	3,3	3,9	2,3	3,1	3,9	4	3
3	3,4	3,1	3,6	2,4	3,1	3,0	4,0	3,0
4	3,7	2,0	4,0	1,7	4,0	2,4	4,0	2,0
5	3,6	2,7	3,4	2,4	2,4	3,0	4	3
6	2,6	2,9	3,4	3,1	3,1	3,0	4,0	3,0
7	2,9	1,6	3,1	1,7	2,6	1,9	3,0	2,0
8	2,7	3,3	2,4	3,7	2,4	3,3	2,0	3,0
M.	3,2	2,5	3,4	2,3	3,0	2,7	3,6	2,8

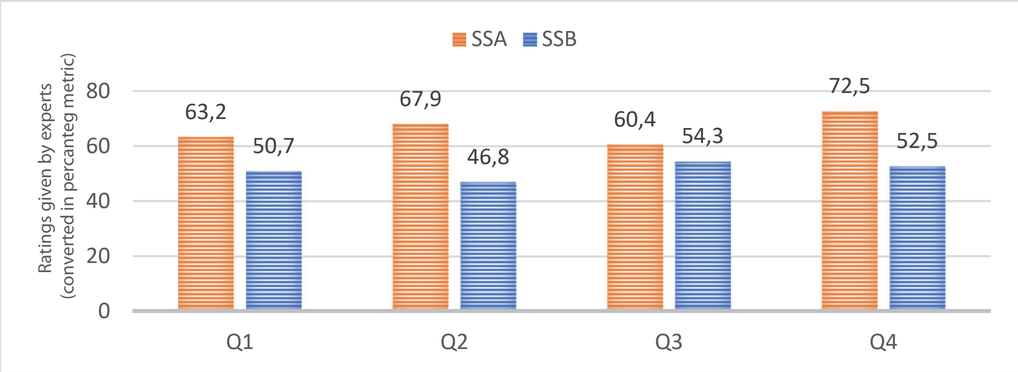


Figure 2
Line chart graph displaying the assessment results over a percentage (%100) metric. (Q: Questions, SSA: Solution Space A -student works, SSB: Solution Space B- DM works)

database of the DM. Since the database of the Midjourney model does not solely contain the image references from basic design education setups, the materialistic aspects of the design solutions were represented in a deflected manner.

The Impact of Ill-Defined Problems

The results of the third question indicate a negligible difference of 6% regarding the performance of solution spaces under ill-defined problems. This result can be interpreted as showing that DMs are as successful as the novice designers in representing the visual implications of ill-defined problems on the generated solutions in substantially faster times. Moreover, the design solutions of the SSA were the best cases among a wide range of students' design works selected for the archive by the instructors. Therefore, if the scope of the study broadens into evaluating the performance of the moderate and mediocre student works, the results would be presumably different in the favor of SSB.

Another consequential point to consider is the significant difference of the generation procedures of SSA and SSB. The SSA instances were generated by the students during the basic design studio procedure, therefore the guide to design solutions was not solely the assignment briefs but also the instructor critiques.

The continuous feedback sessions during the intermediate stages had helped students to comprehend better the implicit requirements of the briefs, as in a way of reflective practice (Schön, 1992b). The abstract concepts, ill-definitions stated in the briefs were discussed repeatedly during the semester, to elucidate the key principles and concepts of basic design. On the other hand, the DM model was only tested by prompting the assignment briefs without any reinforcement learning layer by design expert feedback.

Lost in Process

The interconnectedness of the solution spaces is assessed in the fourth question. The value for SSB during the process is remarkably lower than that for the SSA. This indicates that the evaluated DM is not yet capable of performing a continuous workflow, even if the connections between the processes are given to DM by image-prompting the previous output as a reference input. This situation might be explained by a lack of reasoning mechanisms or a shallow learning curve in the current state-of-the-art DM models compared to the students' reasoning skills in the basic design studio.

This underlines that the implementation of the AI tools in a formal setup of the basic studio would not disregard the importance of the design reasoning mechanisms conveyed at the studio. However, DM models have a potential to be aids in elucidating the design problems for the novice designers.

CONCLUSION

The study demonstrated that the implementation of DMs as a tool for the formal education setup of basic design studio might elevate the comprehension level of novice designers in understanding the implicit definitions of design problems. By generating synthetic solution instances with the help of AI tool the design solution space of the basic design student can further expanded. Since the novice designer has no previous experience in dealing with ill-defined design problems, seeing a myriad of solution alternatives in this expanded design solution space might activate the inherent see-move-see process of design solution generation. Therefore, they can visualize and interpret the defined brief before they start their organic solution delivery process. This could improve their learning curve, and they could easily understand the value and ways of developing a design reasoning mechanism conveyed implicitly in the basic design studio. Furthermore, with a conscious selection of AI generated instances for the panel discussions by their design studio instructors, students would better understand what they need to consider or avoid during the solution generation.

Within research of a broader scope, it would be possible to resolve the previously mentioned two limitations of the study: the context-free training dataset of the DM model; the lack of a feedback process in the generation mechanism. Firstly, it would be possible to generate solutions that fit better into the educational context of a design studio. This can be achieved by training an opensource DM (i.e., Stable Diffusion) with an adequate amount of context-specific data. Secondly, the feedback process with design experts can be incorporated into the generation method of the synthetic design solution space, using reinforcement learning.

The initial DM solutions generated under the sole guidance of the assignment briefs would be revised with an additional set of prompts translated from the design expert feedback. Thus, it would allude to the conventional studio procedure where the design solutions developed with the continuous feedback of the studio instructors.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Tuğyan Aytaç Dural and Mine Özkar for allowing access to their personal basic design studio archives of METU ARCH 101 Basic Design Studio 2004-2005 and Nihal Bursa, Selahattin Önür, Erkan Gencol for their consents. Secondly, we would like to thank all the independent reviewers for contributing to the assessment procedures as design experts.

REFERENCES

Coyne, R. (2005) 'Wicked problems revisited', *Design Studies*, 26(1), pp. 5–17. Available at: https://doi.org/10.1016/j.destud.2004.06.005

Cross, N. (1984). *Developments in Design Methodology*. Chichester: Wiley.

Ghom, P.V., and A. George (2020).' Pedagogy for Basic Design Studio in Learning Architecture: A Qualitative Exploration'. *Journal of Education and Practice,* 11(3). Available at: https://doi.org/10.7176/jep/11-3-18

Midjourney Research Lab (2023). Midjourney Docs.[Online]. Available at: https://docs.midjourney.com/ (Accessed 10 February 2023)

Nanda, S.D., Gagana, G.S., Changela, H.L. and Kumar, N.S. (2021).' Story Visualization: Generation of Scenes sequentially for a given fable using NLP and Image Processing.' *In 2021 Asian Conference on Innovation in Technology (ASIANCON), PUNE, India*. Available at: https://doi.org/10.1109/ASIANCON51346.2021.9545002

Open AI (2023). Dall-E 2. [Online]. Available at: https://openai.com/product/dall-e-2 (Accessed from 10 February 2023)

Saranlı, T. (1998). 'Başlangıçtan Bugüne Temel Tasarım'. In Teymur N.and Aytaç Dural T. (Eds.), *Temel Tasarım / Temel Eğitim* (pp. 37–47). Ankara: ODTÜ Mimarlık Fakültesi Yayınları

Schmidhuber, J. (2015) ‘Deep learning in neural networks: An overview’, Neural Networks, 61, pp. 85–117. Available at: https://doi.org/10.1016/j.neunet.2014.09.003
Schön, D.A(1992a).’ Designing as reflective conversation with the materials of a design situation’. *Research in Engineering Design.* 3. pp. 131-147. Available at : https://doi.org/10.1007/BF01580516
Schön, D.A. (1992b). *The Reflective Practitioner: How Professionals Think in Action (1st ed.)*. Routledge [online]. Available at: https://doi.org/10.4324/9781315237473
Simon, A.H. (2019).’The sciences of the artificial’ in Laird J.(ed).Third edition. Cambridge MA: The MIT Press
Stability AI (2023). Stable Diffusion. [Online]. Available at: https://stability.ai/blog/stable-diffusion-announcement (Accessed from 10 February 2023)
Torfi, A. et al. (2020) ‘Natural Language Processing Advancements By Deep Learning: A Survey’, *ArXiv*, Available at: https://doi.org/10.48550/arXiv.2003.01200
Ustaomeroglu, A. A. et al. (2015) ‘The Impact of Basic Design Studio Courses on Interior Design: KTU Model’, *Procedia - Social and Behavioral Sciences*, 197. Available at: https://doi.org/10.1016/j.sbspro.2015.07.251
Vaswani, A. et al. (2017) ‘Attention Is All You Need’, arXiv [cs.CL]. Available at: http://arxiv.org/abs/1706.03762.
Zbinden, R. (2022). Implementing and Experimenting with Diffusion Models for Text-to-Image Generation. *ArXiv*. Available at: https://10.48550/arXiv.2209.10948