Training, Education, and Learning Sciences, Vol. 155, 2024, 205–215
https://doi.org/10.54941/ahfe1005405

The Impact of AI Image Generator on Traditional Visual Learning: New Possibilities in English Language Learning

Shuo-Fang Liu, Zhi-Cheng Chen, Yi-Chieh Wu, and An-Yu Su

Department of Industrial Design, National Cheng Kung University, Tainan, 701401, Taiwan

ABSTRACT

English learning has always been a crucial topic in the global education sector. The advancement of technology has not only transformed the methods of learning but has also introduced new possibilities for English language acquisition. Traditionally, using images to enhance learners' memory of vocabulary has been an effective method. However, for some abstract words, conveying semantics remains a challenge. With the emergence of AI technology, particularly the application of AI drawing tools, we now have an innovative way to present words. It is anticipated that this approach will have a positive impact on English learning. This study is expected to adopt the concept mapping method combined with the affinity diagram approach. It will be systematically organized and analysed to select five abstract vocabulary terms related to English proficiency tests. Subsequently, AI drawing tools will be employed to generate corresponding images. Following this, a semi-structured interview approach will be utilized to gain an in-depth understanding of users' perspectives on the application of AI drawing tools in English learning and their actual impact on English learning outcomes. The goal of this research is not only to examine the possibility of AI drawing technology replacing traditional images to aid in the memorization of abstract words but also to focus on users' subjective feelings and opinions. We anticipate that the results of this study will provide more innovative and practical approaches to the field of English learning while enriching and expanding existing learning concepts.

Keywords: Image generation, English learning, Concept map, Artificial intelligence, Learning effectiveness assessment

INTRODUCTION

English is one of the most widely used languages in the world, playing a significant role in various fields such as the multimedia industry, business, education, medicine, and engineering (Ilyosovna, 2020). Second language (L2) English learners primarily use English as a tool for communication and interaction with other countries (Sung, 2013). English learning consists of three elements: pronunciation, vocabulary, and grammar. Among these three elements, vocabulary is the most fundamental part (Bai, 2018), and it is crucial for communication, reading, thinking, and learning (Luckner & Cooke, 2010). However, rote memorization in learning may have a counterproductive effect on acquiring English vocabulary (Smith et al., 2013). The research

suggests that images can be considered as one of the strategies for vocabulary learning (Shadiev et al., 2020). This learning method is based on the dual coding theory, establishing connections between language and visuals to enhance information retention (D'Agostino et al., 1977). Therefore, images can be used as visual aids for vocabulary learning in the field of education.

Concrete word is more easily represented by images compared to abstract word (Vellutino & Scanlon, 1985). In recent years, the creativity and efficiency of artificial intelligence (AI) in drawing have been widely acknowledged and studied (Xu et al., 2023). The advantage of AI image generation technology lies in the rapid generation of images (Lyu et al., 2022) and achieving neural style transfer (NST) with the use of Convolutional Neural Networks (CNN) (Jing et al., 2019). AI image generation tools can provide reasonable responses for existing concrete concepts within established training frameworks. However, generating images for abstract concepts requires a higher level of cognition and more interconnected representations. The technology of AI in generating images is still in the exploratory stage, especially regarding abstract vocabulary primarily derived from linguistic experiences.

Based on the foundation that images can aid language learning, this study focuses on

(1) Evaluate users' perceptions of AI-generated images.
(2) Assess the potential for AI-generated images to replace traditional image functions.
(3) Explore the application and impact of AI-generated images on English learning.

LITERATURE REVIEW

Having excellent English proficiency is crucial for many individuals. In recent years, with the rapid development of wireless networks and mobile devices, learning English in a mobile learning environment has become increasingly popular (Chen et al., 2019). Duolingo app features a streamlined gamified interface and breaks down learning into small units, reducing the burden of learning. Voicetube creates an authentic native language environment to enhance listening, speaking, reading, and writing skills. On the other hands, Quizlet utilizes flashcards to improve learning efficiency. Artificial intelligence, as one of the hottest topics in the past three years, has also prompted people to explore its potential applications in foreign language learning. ChatGPT is a chatbot developed by OpenAI, allowing users to obtain information through conversation. Shaikh et al. (2023) conducted usability assessments for English learning in terms of dialogue, writing, grammar, and vocabulary. The research found that ChatGPT has the potential to assist language learning. Based on the strategy of using images for English vocabulary learning, this study focuses on the impact of applying AI technology in generating images on the education sector and the potential for enhancing English learning.

The most common method for vocabulary learning is through the use of pictures. In 1971, Paivio proposed the Dual Coding Theory (DCT), suggesting that the non-linguistic areas of the brain process visual information, while linguistic areas process language information. Throughout this process,

various sensory stimuli are encountered, ultimately establishing connections along psychological pathways (Paivio, 2013). However, abstract vocabulary lacks tangible objects for perception, making it more linguistically complex and challenging to differentiate. Definitions often rely on other vocabulary terms. Due to these characteristics, learners typically need to invest more time and effort in mastering abstract vocabulary, and successful usage requires a high level of language proficiency (Vellutino and Scanlon, 1985). Visual representations of abstract vocabulary can be presented in various ways, including linking oral and visual symbols, connecting through mental images or drawings, and visualizing in specific locations. Association with visual symbols of concrete objects is also possible for abstract words, for example, associating "evil" with the image of a bloody knife (Oxford & Crookall, 1990).

THE EXPERIMENT

The research methodology includes concept mapping, affinity diagram, resume questionnaire and semi-structured interview. Initially, expert panel utilizes the concept mapping method in conjunction with the affinity diagram to select five abstract vocabulary terms suitable for visual representation. Subsequently, images are selected from the internet and generated by an AI image-generating tool, resulting in a total of 10 images—5 from online sources and 5 generated by AI. Each English vocabulary term is associated with one image from the internet and one AI-generated image.

10 participants are then invited to participate in the experiment and interviews. The experiment lasts approximately 30 minutes, consisting of a 5-minute experiment introduction, a 10-minute formal test, and a 15-minute interview (Figure 1). During the experimental process, visual cards and word cards will be utilized as presentation materials. The images generated by AI and selected from the internet will be assigned random numbers to avoid creating Observer-expectancy effect (Rosenthal & Rubin, 1978).

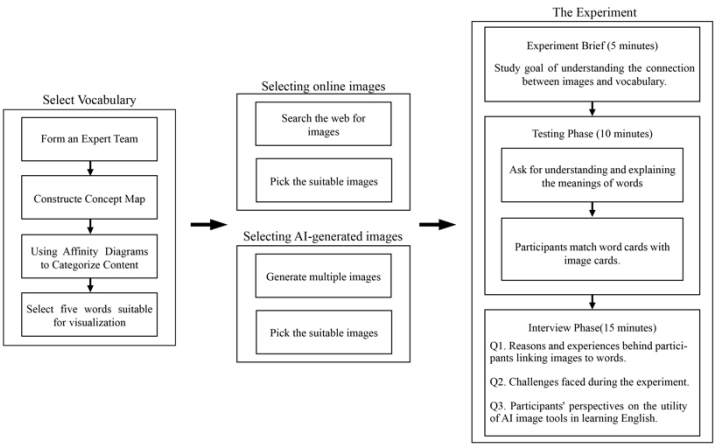

Figure 1: Experimental flow chart.

Concept mapping is a visual graphic tool used to aid in the understanding and organization of information. Concept mapping is a visual graphic tool used to aid in understanding and organizing information. This method was introduced by Novak and Gowin (1984) based on Ausubel's assimilation theory. Concept mapping provide a foundation for understanding, analyzing, and explaining information, and assist researchers making decisions in measurement development and assessment (Rosas & Ridings, 2017). Affinity Diagram, also known as KJ method, is a research method proposed by Jiro Kawakita in 1953. It involves organizing and categorizing research data or various ideas based on their affinity. It is often combined with brainstorming and serves as a quick means of organizing information.

Image tools are divided into internet-based image tool and AI image generation tool. Internet images are sourced from Pinterest website, and AI image generation are operated on Leonardo. Ai. The expert panel selects multiple images for each abstract term from Pinterest and generates 10–15 images by Leonardo. Ai. The most fitting image for the concept of each term is chosen through voting by the expert panel. The technology of Leonardo. Ai is based on the open-source Stable Diffusion. Its advantages include low hardware dependence and the ability to generate up to 150 images for free daily. The operation is similar to Stable Diffusion, involving inputting prompts, providing image input, and applying existing models to generate styles.

Participant

The participants consist of 10 college students, including undergraduates, master's and doctoral students. A resume questionnaire collect anonymous demographic data, including age, educational background, duration of exposure to English, frequency of English usage, and the mediums through which they come in contact with English.

Expert Panel

The expert panel consists of three members: one expert in the integration of technology with English learning, one expert in English learning strategies, and one expert knowledgeable in the operation of concept mapping and affinity diagram processes. All three experts have a minimum of two years of experience using AI tools to support various domains.

Materials

The study is divided into two parts: the first part is selecting five abstract English words, and the second part is choosing 10 images as experimental tools.

Select Vocabulary and Image

In this study, the expert panel utilized concept mapping and affinity diagram methods (Figure 2) to establish the correlation between the characteristics and importance ranking of TOEIC abstract vocabulary. Finally, we identified

and organized 5 abstract vocabulary terms suitable for visualization. "Essential Words for the New TOEIC Test" serves as a reference tool for filtering English vocabulary in this study.

Firstly, the book categorizes proficiency into Basic and High-level sections. In the high-level section's A-Z vocabulary, the expert panel selected words with a 5-star question frequency, using adjectives as the selection criteria. Up to this step, a total of 54 abstract words were selected. Next, the expert panel organized the definitions of abstract words, categorizing them into Nature (characteristics), Form (concrete things, shapes), Relationship (state), Function (concept), and Sensation (feeling). Based on emotional adjectives, sentences tend to be dominantly processed, generating positive or negative evaluations (Lüdtke & Jacobs, 2015). Sensation (feeling) were chosen as the final criterion for the five definitions. Among the 14 selected vocabulary words, a distinction was made between positive and negative vocabulary words. 3 positive vocabulary and 11 negative vocabulary were identified. Ultimately, the expert panel decided to select 5 negative terms from 11 negative words. These terms include: "Morbid", "Vapid", "Aggressive", "Sacrilegious", and "Wanton".

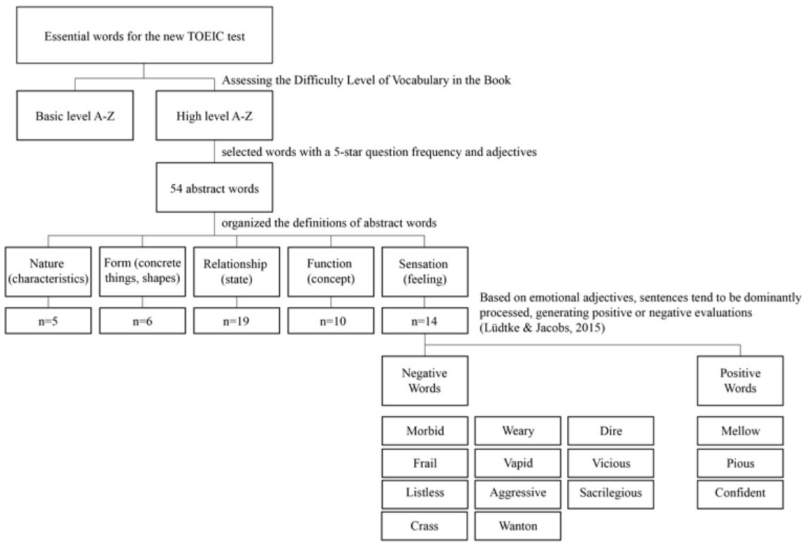

Figure 2: Combining concept mapping and affinity diagram.

For each abstract term, multiple images were collected from the Pinterest image repository. For example, entered "Morbid" into the search bar, and the image repository provided various images related to the word. The expert panel then engaged in discussions to select the most fitting image for each word. To generate AI-generated images, the process involved entering "morbid" in the prompt bar. If satisfied with the generated image, the team utilized the "image-to-image" feature to create images that more closely aligned with the intended meaning team applied the definition or example sentences of "morbid" from online dictionaries to the prompt input field, attempting to

generate images that met the expectations of the expert panel. The final generated images are numbered from 1 to 10, (Table 1) represents the ten images provided to the participants for the assessment.

Table 1. The final images are numbered from 1 to 10.

Words	Morbid	Vapid	Aggressive	Sacrilegious	Wanton
images from the internet	no.10 	no.3 	no.7 	no.8 	no.6 
AI-generated images	no.4 	no.5 	no.2 	no.1 	no.9 

Experimental Procedure

Experiment Brief (5 minutes): Participants received a brief, five-minute explanation outlining the research goal of understanding the connection between images and vocabulary. Consent for interview recording was sought to ensure accurate documentation of responses and context.

Testing Phase (10 minutes): Initially, participants were required to comprehend the meanings of five English abstract words. Researchers queried participants on their understanding of each word, providing explanations for unfamiliar terms. Ten images, labelled 1 to 10, were arranged vertically, and participants matched words to corresponding images. Each word had three corresponding word cards for participants to arrange. The quantity of word cards and image cards wasn't disclosed during the process. The phase concluded at participants' discretion. Researchers later revealed the correct word-image associations to assess accuracy and testing process.

Interview Phase (15 minutes): use semi-structured interviews. Questions centred on participants' experiences, encountered challenges, and perceived learning impacts. The main questions were categorized into three parts:

Q1. Reasons and experiences behind participants linking images to words.

Q2. Challenges faced during the experiment.

Q3. Participants' perspectives on the utility of AI image tools in learning English.

RESULTS

Consolidate and analyze the collected data to obtain information about participants, their performance in the test, and their perceived levels of

association between images and words. Most participants are in the age range of 20–25 years, enrolled in master's programs. The gender distribution is equal, with an even split between males and females. They frequently engage with English through social media and streaming platforms, with an average frequency of 2–3 times per week. Moreover, most participants have had more than three years of contact time.

Data Analysis

Table 2 shows the accuracy rates between images and words. The accuracy rates for no. 2, no. 3, and no. 7 are 10 out of 10, with a highest accuracy rate of 100%. For no. 4 and no. 10, the accuracy rates are 4 out of 10, resulting in the lowest accuracy rate of 40%. Regarding the AI-generated image recognition, the accuracy rates for no. 3 and no. 7 are 10 out of 10, achieving the highest recognition rate of 100%. No. 6 has the lowest recognition rate with 2 out of 10, resulting in a recognition rate of 20%.

Table 2. The participants answering status.

Answer status	Number	Number of people	Correct answer rate
Highest accuracy rate	2, 3, 7	10/10	100%
Lowest accuracy rate	4, 10	4/10	40%
Highest image recognition rate	3, 7	10/10	100%
Lowest image recognition rate	6	2/10	20%

Participants considered images with No. 2, No. 3, and No. 7 to have high relevance, with 10, 9, and 10 respondents respectively. No participant regarded any image as having low relevance. However, in the process of determining relevance, participants identified a middle area where they could understand the meaning conveyed by the image but didn't generate the thoughts about it. Therefore, when exploring images with perceived low relevance, it is not appropriate to solely rely on the quantity of participants. Participants suggested that No. 6, No. 8, and No. 10 were either highly or lowly relevant, lacking a middle ground. It indicates that if participants, upon initial analysis of the images, did not immediately establish a connection between vocabulary and the images, they were unable to comprehend the meaning conveyed by the images.

Table 3. Participants perceived image/word association matching.

Relevance degree	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10
High relevance	5	10	9	2	6	7	10	7	5	5
Low relevance	2	0	0	3	1	3	0	3	2	5

Explanation of Vocabulary and Image Pairings

During the interviews, it was observed that participants provided similar keywords when expressing their impressions of the images. For example,

common keywords for no. 2 include "fierce expression or anger," for no. 6, it is "opening the refrigerator, desire for food," and for no. 8, it is "nun wearing stockings."

There are two controversial images, namely no. 4 and no. 10. Although only 3 participants considered no. 4 to have low relevance, and only 2 considered it to have high relevance, the rest were placed in the middle range. The reason is that no. 4 presented multiple possible corresponding keywords for participants, with key terms like "pale-skinned woman, expressionless face, dark circles under the eyes." These keywords have associations with the meanings of "morbid" and "vapid," making it challenging for participants to make a clear classification.

Users' Perception of AI-Generated Images

Most of the participants have a background in the design field, providing them with a deeper understanding of the operation and artistic techniques involved in AI image generation. And they have applied this knowledge practically in assignments and design competitions.

In terms of the perception of AI-generated images, p01 notes that AI-generated images often exhibit blurred backgrounds, strong lighting changes, and vivid colours, creating a rich visual experience. On the other hand, p03 analyses from personal experience that AI tools tend to have predefined styles, and there is a high probability of generating images that do not match real-world scenes. P10 also acknowledges the richness in the presentation of AI-generated images but emphasizes that richness does not necessarily equate to a good visual presentation.

Possibility of Replacing Traditional Images

In terms of the test content, none of the participants correctly identified 5 AI-generated images and 5 internet-sourced images. This suggests that participants did not specifically perceive a difference between AI-generated and internet-sourced images, and it didn't impact their cognitive judgment of the images.

Some participants think that AI image generation tools are helpful in replacing traditional images, primarily because AI tools have the advantage of quickly generating images, potentially replacing the cost-intensive methods of hand-drawing, photography, or image synthesis (p05–07). while p10 consider that it cannot completely replace traditional images. The reason lies in the fact that traditional image presentation involves human operation. People tend to consider the characteristics of the vocabulary and magnify the impact of that vocabulary in the visual representation, avoiding the occurrence of similar meanings. This helps English learners establish connections between images and words.

Impact of AI-Generated Images on English Learning

In the view of using AI-generated images to assist English vocabulary learning, p09 believes that current AI generation tools still struggle to cleanly distinguish the similarities between words. They can only synthesize data

based on databases, and there is still a gap in terms of image expression compared to traditional images.

The abstract content is derived from external contextual elements such as the environment, objects, physical entities, physical actions, and physical outcomes. Additionally, internal contextual elements including self-relevance, emotions, and motivation also play a role in processing abstract content (Barsalou et al., 2018). Most participants tend to shape their internal contexts based on their personal experiences and external interactions with various elements. p01 expresses that abstract words can be categorized into those that are easily expressed and those that are challenging to express. For instance, words like "vapid" and "aggressive" are more easily linked to personal experiences, making them feel concrete despite being abstract. However, words like "morbid" or "sacrilegious" are difficult to imagine because they cannot be easily connected to scenarios in the mind. The challenge lies in the inability to establish a clear mental context for such words. p08 mentioned that in Japanese anime, there is often a character known as "yandere" who develops intense feelings for a person or thing, to the point of approaching madness or illness. During the test, they could quickly associate "morbid" with the two controversial images, no. 4 and no. 10. This ability is attributed to their personal experience of watching anime, which projects highly relevant images associated with "morbid" in their minds.

DISCUSSION

Users' perspectives on the use of AI-generated images for English learning provide profound and diverse insights. AI image generation tools have advantages in enriching visual representations and cost-effectiveness. However, careful evaluation of their limitations is needed when it comes to replacing traditional images and assisting in English vocabulary learning. However, abstract nouns rely heavily on language systems from a neurobiological perspective (Sabsevitz et al., 2005), and AI image generation tools have not yet established the ability to communicate with humans (Zhang et al., 2023). Traditional image expressions, guided by human operation, exhibit more humanistic thinking and creativity, placing AI image generation tools at a disadvantage. The ease of expressing abstract words is also a factor, as those easily expressed can be concretely presented and universally recognized by AI image generation tools. But challenging abstract words require a combination of external and internal factors to achieve the association between images and vocabulary.

CONCLUSION

This study investigated the impact of AI image generation tools on English vocabulary learning through experiments and interviews. The results indicate that AI-generated images can reduce the time and cost of image production and have the potential to replace traditional images. However, they may not necessarily aid in English vocabulary learning. The reason lies in the distinction between concrete and abstract concepts within English vocabulary. This

study specifically focused on exploring the use of images to assist in learning abstract vocabulary. Abstract words lack concrete reference points, making learning more challenging. The study found that AI image generation technology cannot fully overcome the disadvantages of traditional images, and there are still difficulties in distinguishing between similar words, creating a certain gap compared to traditional image expression.

Through experiments and interviews, it was observed that abstract vocabulary exhibits variations in cognitive expression. However, the technology of AI image generation can only capture the representation of a portion of vocabulary and not comprehensively. The current application of AI-generated images in aiding English learning is limited, but the technology has already been applied to the field of language learning through other means. Optimism is maintained in this study for the potential of AI-generated images to aid in English learning, anticipating further refinement of AI technology in the future.

ACKNOWLEDGMENT

The Consortium is funded by the National of Science and Technology Council (NSTC) (MOST 111-2221-E-006-083-MY2). Thank you all.

REFERENCES

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). Educational psychology: A cognitive view.

Bai, Z. (2018). An analysis of English vocabulary learning strategies. Journal of language Teaching and Research, 9(4), 849–855.

Barsalou, L. W., Dutriaux, L., & Scheepers, C. (2018). Moving beyond the distinction between concrete and abstract concepts. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1752), 20170144.

Chen, C.-M., Chen, L.-C., & Yang, S.-M. (2019). An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. Computer Assisted Language Learning, 32(3), 237–260.

Cheung Matthew Sung, C. (2013). Learning English as an L2 in the global context: Changing English, changing motivation. Changing English, 20(4), 377–387.

D'Agostino, P. R., O'Neill, B. J., & Paivio, A. (1977). Memory for pictures and words as a function of level of processing: Depth or dual coding? Memory & Cognition, 5, 252–256.

lyosovna, N. A. (2020). The importance of English language. International Journal on Orange Technologies, 2(1), 22–24.

Jing, Y., Yang, Y., Feng, Z., Ye, J., Yu, Y., & Song, M. (2019). Neural style transfer: A review. IEEE transactions on visualization and computer graphics, 26(11), 3365–3385.

Luckner, J. L., & Cooke, C. (2010). A summary of the vocabulary research with students who are deaf or hard of hearing. American annals of the deaf, 155(1), 38–67.

Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: additive or interactive effects of nouns and adjectives? Frontiers in psychology, 6, 1137.

Lyu, Y., Wang, X., Lin, R., & Wu, J. (2022). Communication in human–AI co-creation: Perceptual analysis of paintings generated by text-to-image system. Applied Sciences, 12(22), 11312.

Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. cambridge University press.

Oxford, R., & Crookall, D. (1990). Vocabulary learning: A critical analysis of techniques. TESL Canada journal, 09–30.

Paivio, A. (2013). Imagery and verbal processes. Psychology Press.

Paivio, A. (2014). Bilingual dual coding theory and memory. Foundations of bilingual memory, 41–62.

Rosas, S. R., & Ridings, J. W. (2017). The use of concept mapping in measurement development and evaluation: Application and future directions. Evaluation and program planning, 60, 265–276.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. Behavioral and Brain Sciences, 1(3), 377–386.

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. Neuroimage, 27(1), 188–200.

Shadiev, R., Wu, T.-T., & Huang, Y.-M. (2020). Using image-to-text recognition technology to facilitate vocabulary acquisition in authentic contexts. ReCALL, 32(2), 195–212.

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal english language learning. European Journal of Investigation in Health, Psychology and Education, 13(9), 1937–1960.

Smith, G. G., Li, M., Drobisz, J., Park, H.-R., Kim, D., & Smith, S. D. (2013). Play games or study? Computer games in eBooks to learn English vocabulary. Computers & Education, 69, 274–286.

Vellutino, F. R., & Scanlon, D. M. (1985). Free recall of concrete and abstract words in poor and normal readers. Journal of experimental Child psychology, 39(2), 363–380.

Xu, J., Zhang, X., Li, H., Yoo, C., & Pan, Y. (2023). Is Everyone an Artist? A Study on User Experience of AI-Based Painting System. Applied Sciences, 13(11), 6496.

Zhang, C., Wang, W., Pangaro, P., Martelaro, N., & Byrne, D. (2023). Generative Image AI Using Design Sketches as input: Opportunities and Challenges. Proceedings of the 15th Conference on Creativity and Cognition.