Artificial Intelligence Text-to-Image Generation in Art and Design Teacher Education – A Creative Tool or a Hindrance to Future Creativity?

Tore Andre Ringvold, Oslo Metropolitan University
torand@oslomet.no
Ingri Strand, Oslo Metropolitan University
Peter Haakonsen, Oslo Metropolitan University
Kari Saasen Strand, Oslo Metropolitan University

ABSTRACT

In today's constantly changing world technological developments in artificial intelligence (AI) can induce educational visions of both utopia and dystopia. New technologies and communication platforms can provide new forms and possibilities of learning. Creating an image has historically mostly been a human process of using knowledge and application of technique that demanded training. This image-making process changed with the invention, development and spread of the photographic camera, when creating a detailed visual representation of reality became a possibility without a complex process of craftsmanship and artistry. The nature of visual art changed but the visualisation of ideas and prefigurative thoughts could not necessarily be captured by a camera. With the development and spread of AI text-to-image generation, can this change the need for competency to visualise ideas in the way the camera changed the need for drawings and paintings as visual representations?

This study explores how AI text-to-image generators can contribute to and change art and design teacher education. We conducted exploratory experiments where we tested a variety of AI text-to-image generators and explored the outcome of using different generators, prompts and settings. Reflections were written down throughout the process. This was combined with an online ethnography on text-to-image communities. Different potentials of learning were identified, as well as issues of interaction and possible contexts of use. The results are discussed in a future learning context.

Key Words: Artificial intelligence, Text-to-image generation, Imagination, Art and design teacher education, Visualisation

1. INTRODUCTION

Education and the way we learn adapt to society. Constant changes in technology have led to new cultures and places of learning (Thomas & Brown, 2011). Artificial intelligence (AI) represents

new digital technology that has great power to change society. With the rapid development of AI technology, educators need to adapt to emerging technology that can potentially change how we work and produce. Artificial intelligence text-to-image technology has been developed since 2014 and increased in popularity after the introduction of DALL-E in 2021 (Cetinic & She, 2022). In 2022, this development accelerated rapidly with the launch of services such as Midjourney, OpenAI's DALL-E 2, and Stable Diffusion. Since this is such a recent technology, published research on the topic is limited. With the widespread use and rapid technological development, there is a need for research to critically look at the possible use and application of AI text-to-image generation. In this study, we ask: How can AI text-to-image generators contribute to and change art and design teacher education?

2. BACKGROUND

2.1. A short history of visualisation

Prefigurative thinking is seen as an important part of what constitutes being human (Fry, 2012). The prefigurative thought of an idea or concept can be materialised through medium and technique. Doing this and also being good at it are no easy tasks. As most have experienced, it takes time to develop the knowledge and skills to produce traditional visual forms that constitute sensorimotor activities such as drawing and painting. There were no quick-fix methods for creating quality visualisations before the introduction of photography. The process of image creation underwent a significant transformation with the advent, progress, and widespread use of the photographic camera (Gombrich, 1982). It enabled the creation of highly accurate visual depictions of reality without intricate craftsmanship and artistic techniques. However, as prefigurative thoughts cannot necessarily be captured by a camera, there is a need for skills in drawing and illustration (Nielsen, 2013). Can the development and spread of AI text-to-image generation change the need for competencies to visualise ideas in the way the camera changed the need for drawings and paintings as mimetic visual representations?

2.2. Text-to-image generation technology

There are different text-to-image generative systems, but what they have in common is that textual inputs (prompts) are interpreted by a system before images are created. The systems are trained on large datasets of text-and-image pairs from the web (Abdallah & Estevéz, 2023). A prompt can lead to unexpected results, but at the same time the different models, such as Midjourney and OpenAI (DALL-E 2), provide tips on how to alter the style or format by adding specific terms. These tips can help one to affect the outcome and, by adjusting the prompt input, one can increasingly control the image-making process. The images in this paper were mostly created with Midjourney Version 4, which was the default model from November 2022 to May 2023, with Version 5 released on 4 May 2023 (Midjourney, n.d.).

2.3. A changing art and crafts education providing competencies for the future

Norwegian art and design teacher education qualifies students for teaching at all levels, from six-year-old first graders to 18–19-year-old upper secondary school students preparing for work or

further studies. In primary and lower secondary school, Art and crafts is a compulsory subject. The 2020 curriculum revision ensures a more future-directed education, in part by making technological use and programming more visible (Norwegian Directorate for Education and Training, 2020). Art and design teachers are also qualified to teach at several upper secondary school programmes, with art, craft and design constituting a central part of their curricula, either specializing in fields such as traditional craft or digital media production. Current pre-service teachers will teach over the next 40 to 50 years, and the extensive development of AI in recent years renders it impossible to predict what the future will look like so far into the future. Rather than transferring knowledge that they in turn will teach, teacher education must prepare students to facilitate future learning processes regarding future challenges.

Artificial intelligence will challenge schools' teaching and assessment practices, and a new government strategy (*Kunnskapsdepartementet*, 2023) requires schools to adopt AI in order to gain a basic understanding of how AI works, its solutions, and its limitations. Thus, teachers and students need sufficient competencies to use AI in an exploratory way, with curiosity, critical thinking, and ethical awareness.

3. METHODOLOGY

We explored AI text-to-image generators over a six-month period from November 2022 to May 2023. We chose a qualitative approach, combining explorative experiments and online ethnography. Dyrssen (2010) states that, while explorative experiments cannot be validated, they allow the researcher(s) to 'shake up ingrained patterns of thought; provide quick feedback, increased curiosity, and discoveries of hidden possibilities; reveal possible links and points that need to be mapped; and get the creative process moving forward' (p. 229).

Each of the four authors tested AI text-to-image generators and explored the outcome of different generators, prompts and settings in use. The authors had little to no experience with AI text-to-image generators and wrote prompts based on their own imagination and curiosities, sometimes choosing to follow interesting idea strands. The combined variety of prompt inputs, ranging from the abstract to the concrete, is reflected in the examples presented in section 4. Throughout the process, reflections were written down and image material saved. Our explorations coincided with significant developments of the generators, such as the release of Midjourney Version 5.1 and solutions to recurring issues such as the depiction of fingers (Verma, 2023). Due to the rapid development, we are more concerned with the overarching concept of AI text-to-image generators than specific technical aspects.

The authors shared their experiences in frequent meetings, drawing on their own backgrounds in design and education. Each author has approximately 10 years of teaching experience covering all levels, ranging from Year 4 (nine-year-olds) to university level. Through reflective dialogue, we discovered central issues regarding our experiences of AI text-to-image generators, as well as staking out a path for further experiments. The first author also conducted an online ethnography (Hart, 2017; Winter & Lavis, 2020) in the text-to-image community of the Midjourney server on Discord, capturing users' interactions at three different timepoints.

Through a thematic analysis (Braun & Clarke, 2022) of the empirical material from the explorative experiments and the online ethnography, we discovered themes related to AI text-to-image generators' potential for learning, limitations, and hindrances, as well as issues of interaction and possible contexts of use. The results from this analysis are presented and discussed in the following section.

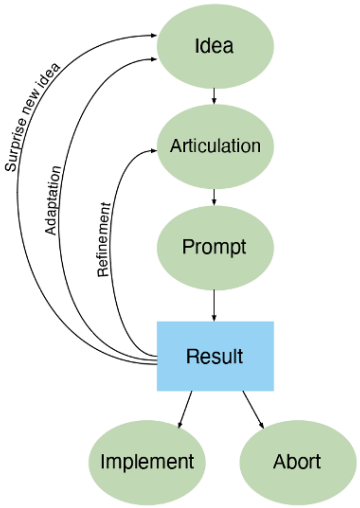
4. A TEACHER TRAINING LEARNING PERSPECTIVE ON AI TEXT-TO-IMAGE TECHNOLOGY

The results and discussion are presented in a bilateral manner of learning potentials on the one side and learning obstacles on the other. However, the identified qualities are not necessarily either a potential or a hindrance to learning; they can be both at the same time.

4.1. Modes of creation processes from a learning perspective

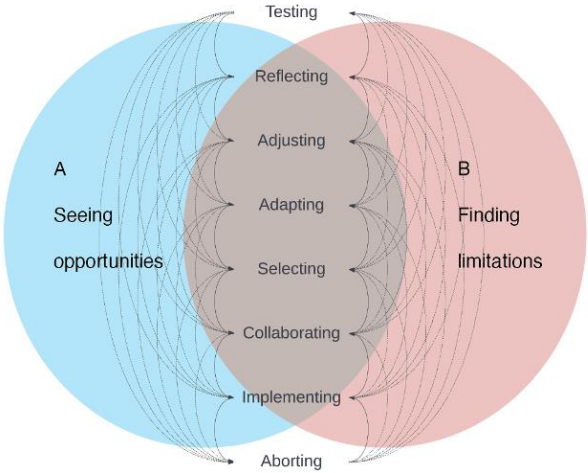
Creating visualisations using text-to-image generators is a cyclic, iterative process. We identified this image creation process as consisting of several stages, as visually presented in Figure 1. The first of these is the idea stage. This can be a prefigurative thought at the beginning of a process, the refinement of previous ideas, an adaptation of a previous idea, or a surprising new idea that originated from a previous process. Taking the step from idea to a written prompt, one needs to shape the visual idea or concept through articulation (stage 2). After feeding the prompt in writing (stage 3), a black box process (Bunge, 1963) gives one the results (stage 4). Results can be implemented (used) or aborted (left unused). They can also be refined, adapted, or used to start new ideas in a cyclic process.

Figure 1.
The identified image creation process using AI text-to-image generation.



In addition, we propose an understanding of the learning process that takes place while using AI text-to-image technology (Figure 2). In a cyclic, interactive learning process between the human user and text-to-image AI technology, a variety of actions can take place in an interchange between A: Seeing opportunities and B: Finding limitations. These actions can contribute to the cyclic development of learning processes of varying length. Variations also apply to the actions involved in the cycle.

Figure 2.
A proposed representation of the cyclic human user and AI text-to-image interactive learning process.



4.2. 4.2. Potential learning opportunities

4.2.1. Play with imagination: A training exercise in prefigurative thinking?

Artificial intelligence text-to-image technology provides the user with a powerful tool to generate visualisations without much effort. This ease of access to visualising allows for playful image-making. The image creation could, especially at the beginning of using the technology, be about testing limits and seeing what is possible, as illustrated in Figure 3. Using this technology, students can possibly push their own boundaries of imagination. The ease of access can trigger a willingness to try and experiment with creative image making. Early adolescents may experience a more critical view of their produced drawings, as described by Lowenfeld (1947). For the non-professional image creator, text-to-image generation can provide a beneficial training ground and a tool for visualising, both mentally and physically. This is especially relevant for those who are hesitant to enter the creative processes of image making.

Figure 3.
Novelty wow factor tests created with Midjourney (V4). Prompt 1 (to the left): Unimaginable thoughts happy extreme. Prompt 2 (to the right): A muppet riding a centaur galloping along a windy beach.

4.2.2. An online arena for sharing and collaborative creativity

Midjourney and Discord provide an arena for knowledge-sharing and collaborative generative art and design processes. These processes take place via a variety of chatrooms, show-and-tell rooms and other channels of communication. In these multi-human and machine collaborations, several users and the AI generative technology are part of discussions. Through discussions and testing, they collaborate on developing prompts, aesthetic qualities and designs. This online space for potential collaboration and co-learning provides a learning environment independent of place. Figures 4.1, 4.2, 4.3 and 4.4 provide examples of Midjourney users collaborating and helping each other to achieve a desired image. Collaboration and seeking guidance from others through sharing is an important part of creative professional work and are also highlighted in new curricula in primary and secondary education (Norwegian Directorate for Education and Training, 2020).

Figure 4.1.
Attempting a human-like Marge Simpson through prompting. Chat conversation, requesting prompt help. Discord, Midjourney server, prompt-chat screenshot.

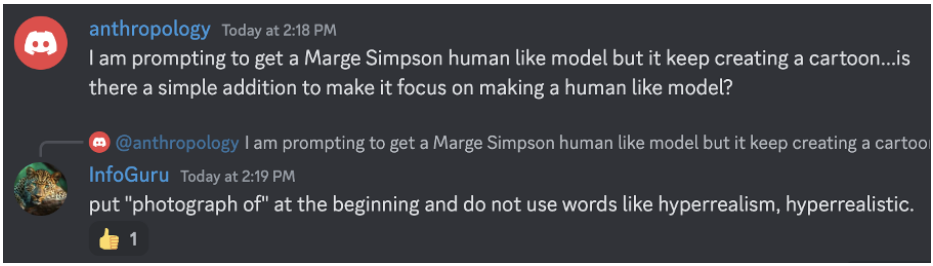


Figure 4.2.
Help suggested. Discord, Midjourney server, prompt-chat screenshot.

Figure 4.3.
Help suggested, follow-up. Discord, Midjourney server, prompt-chat screenshot.

Figure 4.4.
A third user joins the discussion with a comment and showing their own take on the prompt. Discord, Midjourney server, prompt-chat screenshot.

4.2.3. A visual aid for developing perspectives and composition

Artificial intelligence images can be useful visual aids for creative work. In our exploration, one of the authors tried this approach to draw compositions for a comic book, including an image of a large dragon sleeping in a parking house. Using AI, he could explore how to draw this composition. Images created with Midjourney often do not meet the creator's expectations, but the images can be used as sublayers for reference in constructing illustrations. By prompting several iterations of a dragon and the environment and combining multiple images to get a reliable perspective (Figure 5), the drawing could be completed.

Figure 5.
Dragon and environment image generations combined with author's own sketch work (far right), used for reference in creation process.

4.2.4. A multitude of idea generations

The study identified that, in creating object designs, the AI generated a multitude of varying idea visualisations. A series of relevant designs were easily created when the object shape was conventional, such as a ring (Figure 6), or a conventional design such as a car or a toaster (Figure 7). The AI does not rethink a functional need or type of use but generates imagery based on convention. This does not mean that all suggestions are conventional. Mashups of existing design visualisations can provide interesting results or suggestions useful for further exploration. The variations are more decorative rather than shape variations. In a creative learning process, students' abilities to generate ideas differ. A common obstacle in the creative process can be a lack of variation of ideas in the early stages, leading to unsatisfactory results that do not fulfil their potential. The ease of generating idea visualisations can help prevent concept or design fixation (Schut et al., 2020). Artificial intelligence text-to-image technology can provide new avenues for solutions to a problem or new ways of seeing and understanding.

Figure 6.
Different engineering department alumni ring design propositions created using Midjourney (V4). Prompts: An engineering signet ring 21st century, construction computing, and mathematical.

Figure 7.
Different toaster design propositions created using Midjourney (V5). Prompts: A toaster made of delicate porcelain, sheer origami paper, mosaic.

4.2.5. Potluck visualisation as a tool for creativity

Artificial intelligence text-to-image generation can be a game of chance, providing surprising image results. This potential potluck quality of AI text-to-image generations led us to look at our initial ideas in new and unforeseen ways. Opening new avenues of thought can be rewarding in the creative process. While providing new possibilities, it also emphasises the need for a more critical creator role, as described in section 4.2.7. Using chance or coincidence in idea generation can contribute to new perspectives that enrich the creative process (Fazel & Almousa, 2021). Surprising image results, unrelated to the original idea, can be exciting to develop further. In these cases, the process, or 'dialogue' with the text-to-image generator, produces new and unexpected outcomes. This resonates with Oppenlaender (2022, p. 198), who has referred to unexpected results and serendipity.

4.2.6. Developing and strengthening a visual reference language

We identified articulation as an important stage between idea and prompting (Figure 1). The training of articulation from prefigurative thought to writing can contribute to the development and strengthening of a visual reference language. On the one hand, the ability to use the subjects' terminology is often highlighted in curricula. On the other hand, when articulating and writing prompts, one might use words or phrases not necessarily viewed as correct in the art and design discipline. However, they work well when prompting due to their widespread use online, for example in gaming communities.

4.2.7. The 'concept-articulating catalyst wizard', a changing role of the image creator?

The study's exploration identified that the role of creator in creative processes using AI text-to-image technology differs from the creator's role in traditional image making such as drawing. We see the creator's role shifting towards that of an art director, composer, editor, or selector. The process role of editing and catalysing black box processes (Oppenlaender, 2022; Vartiainen & Tedre, 2023) consists of articulating, testing, developing, adapting, refining, selecting and editing by starting new cyclic processes, as described in Figures 1 and 2. This shift in the creator's role questions what kind of knowledge and skills will be needed in future creative processes. With a shift towards editing and selecting rather than producing, a critical mindset should be an important part of future creator and design competencies. Such a mindset should be critical of results and open to different solutions of visualising prefigurative thoughts.

4.3. Potential learning obstacles

4.3.1. Chasing the centaur, not getting what one wants

An example from our testing was to create a centaur by using different prompt-writing approaches. Whether writing a short prompt, such as 'centaur', or describing what a centaur is, the results were mainly images of horses. Other results depicted a man standing in front of a horse-like body, or a human torso attached to a horse's back. Although other Discord users in the Midjourney community had managed to create centaurs, a successful prompt copied and pasted from the community also elicited poor results. Other absurd combinations were also difficult to accomplish. The first attempts at creating AI images may be fascinating, but the wow factor will not necessarily last for long. Trying to create something based on ideas and imagination may lead to disappointing results that do not match how one visualised the ideas in the first place. The natural limitations of a given technology or tool will limit the possibilities.

4.3.2. Ethics, copyright and censorship

Who owns images created with AI? Being trained on large datasets from the internet, there is no guarantee that images will not violate any copyrights (Abdallah & Estévez, 2023). In a classroom setting one can explore and experiment with images without violating copyrights (Bergman, 2021). However, as Midjourney shares the generated images with the community and the creator alike, one can question which copyrights are potentially violated with each image generation. Making images in an AI-based process provides a natural ground for discussing ethical dilemmas of copyright infringements and obstacles of censorship. Due to different types of censorship, some of the AI models have certain constraints. With Midjourney, one cannot create imagery based on

prompts that suggest sexual content, while materials potentially violating copyright or personal data issues seem to be accepted. Compared to Midjourney (Figure 7), the new beta version of Adobe Firefly seems more restrictive. These limitations in technology due to censorship, copyright or privacy issues limit the user's freedom to express their prefigurative thoughts in visualisations.

Figure 8.
A comparison of two different providers with the same prompt, Realistic photo Mickey Mouse portrait. Left: Midjourney V5. Right: Adobe Firefly (beta).

4.3.3. Bias and stereotypes

As the images generated by text prompts are based on image-and-text pairs from the internet, biases and stereotypes may be reproduced (Vartiainen & Tedre, 2023). If ethnicity or gender is not specified in a prompt, a white male is often featured in the results, unless the prompt describes people not associated with men, for example, a nun or a geisha. In an attempt to create images of a female adult piano player, Midjourney generated images of skinny young women and girls, although the prompt contained words such as '40-year-old', '50-year-old', and 'middle-aged'. Finally, after asking for a 'geriatric' woman, the results did not include young girls.

4.3.4. A cop-out?

The low-effort ease of image making with AI makes us question whether its extensive use can result in a non-critical view on the benefits of making. If the use of AI text-to-image technology replaces sensorimotor making activities such as drawing, what is possibly lost in a creation process consisting less of producing and more of articulating and selecting? In traditional image-making processes consisting of applying sensorimotor techniques, training is essential. Drawing skills are developed over multiple years of practice. If such sensorimotor skills for producing imagery based on prefigurative thinking become superfluous, what is the incentive in education to develop sensorimotor skills for mimetic drawing, as described by Nielsen (2013)? Is this ease

of use a hindrance to learning or can it possibly free up more time to focus on other knowledges and skills necessary in a future learning environment? The use of AI technology in learning processes demands a critical teaching mindset, ensuring the necessary training and development of the skills needed.

5. CONCLUDING REMARKS

Artificial intelligence text-to-image technology can contribute to and change future art and design education in various ways. It can contribute to increased opportunities for training prefigurative thinking, providing new ways of visualising and co-creating. This can represent a democratization of visualising prefigurative thinking, as creators are not being restricted by their limitations in skills or techniques. The use of AI can simplify and enrich image making and design processes. Artificial intelligence text-to-image technology can represent a useful tool for creative processes and developing articulation for visualising the imaginary. Its use can also represent limitations to creativity and contribute to ethical questions and issues of bias being raised. With the application and use of AI text-to-image technology in art and design education spreading, we need to question what kinds of competencies are needed in future learning processes.

This limited study has identified several possible avenues for further research. With the widespread use and development of this technology, we emphasize the need for a critical perspective in future research.

6. REFERENCES

Abdallah, Y. K., & Estévez, A. T. (2023). Biomaterials Research-Driven Design Visualized by AI Text-Prompt-Generated Images. *Designs*, 7(2), 48. MDPI AG. http://dx.doi.org/10.3390/designs7020048

Bergman, S. (2021). Copying and cultural references as tools in image creation: Copyright challenges in teaching. *Nordic Journal of Art & Research*, 10(1). https://doi.org/10.7577/information.3692

Braun, V., & Clarke, V. (2022). Thematic analysis: a practical guide. SAGE.

Bunge, M. (1963). A General Black Box Theory. *Philosophy of Science*, 30(4), 346–358. http://www.jstor.org/stable/186066

Cetinic, E. & She, J. (2022). Understanding and Creating Art with AI: Review and Outlook. *ACM Transactions on Multimedia Computing, Communications and Applications*, 18(2), Article 66. https://doi.org/10.1145/3475799

Dyrssen, C. (2010). Navigating in Heterogeneity: Architectural Thinking and Art-Based Research. In M. Biggs, H. Karlsson (Eds.), The Routledge Companion to Research in the Arts (pp. 223–239). Routledge.

Fazel, M., & Almousa, S. (2021). Chance, Ambiguity, and Indeterminacy as Idea-Generating Mediums Applied in Creative Design: Encountering Uncertainty in Mediums of Drawing in the Design Process. In F. Bianconi & M. Filippucci (Eds.), Digital Draw Connections: Representing

Complexity and Contradiction in Landscape (pp. 267–293). Springer International Publishing. https://doi.org/10.1007/978-3-030-59743-6_12

Fry, T. (2012). Becoming Human by Design. Bloomsbury Publishing Plc.

Gombrich, E. H. (1982). The Image and the Eye. Phaidon Press Limited.

Hart, T. (2017). Online Ethnography. The International Encyclopedia of Communication Research Methods, 1-8.

Kunnskapsdepartementet [Ministry of Education and Research]. (2023). *Strategi for digital kompetanse og infrastruktur i barnehage og skole*. https://www.regjeringen.no/no/dokumenter/strategi-for-digital-kompetanse-og-infrastruktur-i-barnehage-og-skole/id2972254/

Lowenfeld, V. (1947). Creative and Mental Growth. Macmillan. Retrieved June 7, 2023. https://archive.org/details/creativementalgr1evikt

Midjourney, (n.d.). Model Versions. Retrieved June 13, 2023. https://docs.midjourney.com/docs/model-versions

Nielsen, L. M.(2013) Visualising ideas: a camera is not enough, in Reitan, J. B., Lloyd, P., Bohemia, E., Nielsen, L. M., Digranes, I., & Lutnæs, E. (eds.), DRS // Cumulus: Design Learning for Tomorrow, 14–17 May, Oslo, Norway. https://doi.org/10.21606/learnxdesign.2013.103

Norwegian Directorate for Education and Training. (2020). Curriculum for Art and Crafts (KHV01-02). The National Curriculum *Kunnskapsløftet* 2020. https://www.udir.no/in-english/curricula-in-english/

Oppenlaender, J. (2022). The Creativity of Text-to-Image Generation. In 25th International Academic Mindtrek Conference (Academic Mindtrek 2022), November 16–18, 2022, Tampere, Finland. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3569219.3569352

Schut, A., Klapwijk, R., Gielen, M., van Doorn, F., & de Vries, M. (2020). Uncovering early indicators of fixation during the concept development stage of children's design processes. *International Journal of Technology and Design Education*, 30(5), 951–972. https://doi.org/10.1007/s10798-019-09528-2

Thomas, D., & Brown, J. S. (2011). A New Culture of Learning: Cultivating the Imagination for a World of Constant Change. CreateSpace Independent Publishing Platform.

Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-to-image generative models. Digital Creativity, 34(1), 1-21. https://doi.org/10.1080/14626268.2023.2174557

Verma, P. (2023, 26 March). AI can draw hands now. That's bad news for deep-fakes. *Washington Post*. https://www.washingtonpost.com/technology/2023/03/26/ai-generated-hands-midjourney/

Winter, R., & Lavis, A. (2020). Looking, but not listening? Theorizing the practice and ethics of online ethnography. Journal of Empirical Research on Human Research Ethics, 15(1–2), 55–62.