

# THEMATICPLANE: Bridging Tacit User Intent and Latent Spaces for Image Generation

Daniel Lee

Adobe Inc.

San Jose, CA, USA

dlee1@adobe.com

Nikhil Sharma

Johns Hopkins University

Baltimore, MD, USA

nsharm27@jhu.edu

Donghoon Shin

University of Washington

Seattle, WA, USA

dhoon@uw.edu

DaEun Choi

KAIST

Daejeon, Republic of Korea

daeun.choi@kaist.ac.kr

Harsh Sharma  
University of Colorado  
Boulder, CO, USA  
harsh.sharma@colorado.edu

Jeonghwan Kim  
UIUC  
Champaign, IL, USA  
jk100@illinois.edu

Heng Ji  
UIUC  
Champaign, IL, USA  
hengji@illinois.edu

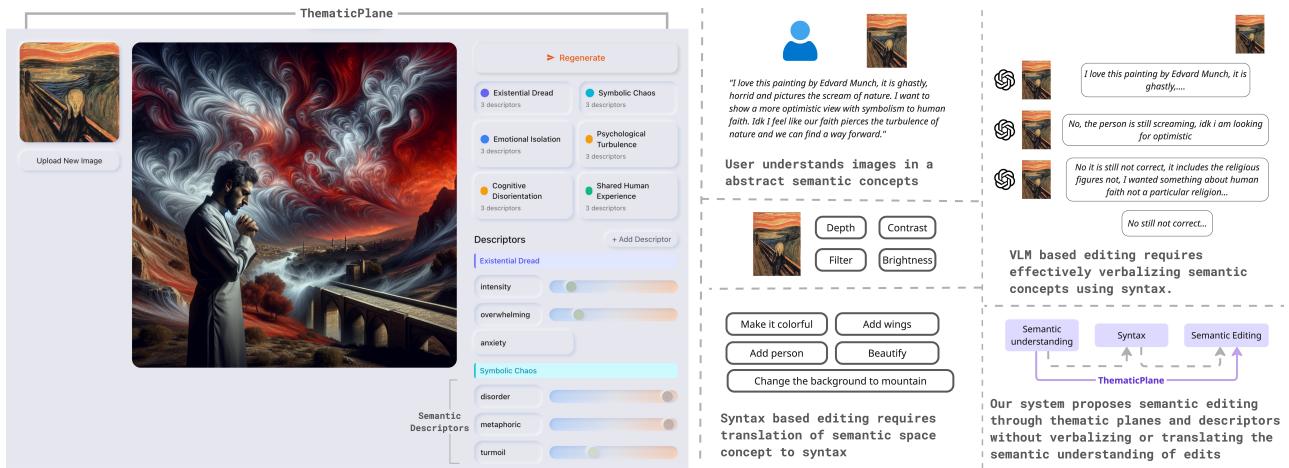


Figure 1: Overview of **THEMATICPLANE**. Unlike conventional prompting, **THEMATICPLANE** supports the navigation and manipulation of high-level semantic concepts within a thematic plan, allowing for tacit editing of the image aligned with the user's understanding of the semantic concept of the image.

## Abstract

Generative AI has made image creation more accessible, yet aligning outputs with nuanced creative intent remains challenging, particularly for non-experts. Existing tools often require users to externalize ideas through prompts or references, limiting fluid exploration. We introduce **THEMATICPLANE**, a system that enables users to navigate and manipulate high-level semantic concepts (e.g., mood, style, or narrative tone) within an interactive thematic design plane. This interface bridges the gap between tacit creative intent and system control. In our exploratory study ( $N = 6$ ), participants engaged in divergent and convergent creative modes, often embracing unexpected results as inspiration or iteration cues. While they grounded their exploration in familiar themes, differing expectations of how

themes mapped to outputs revealed a need for more explainable controls. Overall, **THEMATICPLANE** fosters expressive, iterative workflows and highlights new directions for intuitive, semantics-driven interaction in generative design tools.

## Keywords

creativity support tool, visual exploration, generative AI

## ACM Reference Format:

Daniel Lee, Nikhil Sharma, Donghoon Shin, DaEun Choi, Harsh Sharma, Jeonghwan Kim, and Heng Ji. 2025. **THEMATICPLANE: Bridging Tacit User Intent and Latent Spaces for Image Generation**. In *The 38th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '25)*, September 28–October 1, 2025, Busan, Republic of Korea. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3746058.3758376>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UIST Adjunct '25, Busan, Republic of Korea

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2036-9/2025/09

<https://doi.org/10.1145/3746058.3758376>

## 1 Introduction

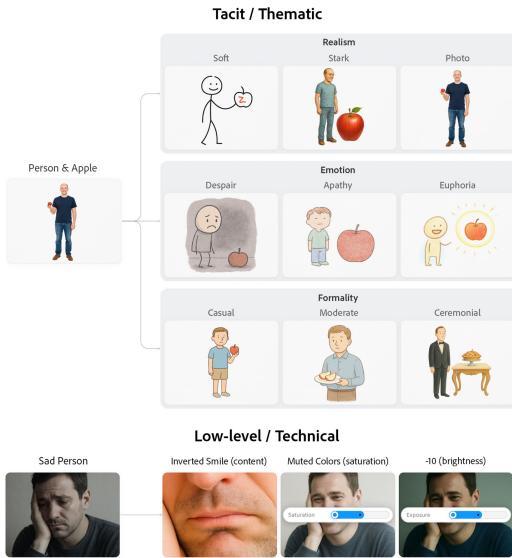
The proliferation of generative AI has fundamentally transformed creative expression in various modalities (e.g., text, images, video, code, etc.), enabling users to synthesize artifacts with unprecedented ease (e.g., via tools like ChatGPT or RunwayML). For instance, individuals without expertise in visual editing software

can now generate or modify images through simple prompts, enabling lay users to construct sophisticated outputs [16]. Nonetheless, achieving precise, intent-aligned results still demands technical knowledge (e.g., manually crafting detailed prompts, understanding the underlying model behaviors), hindering non-expert users [20].

Previous works have proposed systems where users can change low-level image parameters (e.g. with augmented prompt engineering scaffoldings [9, 24]). However, users seldom conceptualize desired changes in terms of such low-level technical parameters, layers, or explicit instructions. Instead, they reason in high-level semantic terms, like aspiring to make an image ‘warmer,’ ‘more dramatic,’ or ‘friendlier’ [4, 7]. This semantic gap as visualized in Figure 2—between users’ tacit, thematic intent and systems’ affordances—has spurred recent HCI research into intuitive editing paradigms, including keyword recombinations [5], reference image-based modifications [21], and interactive latent space explorations [1, 17].

Despite these, existing tools predominantly require users to externalize their vision through prescribed mediums, such as textual descriptions or reference images. This ‘externalization assumption’ presumes users can readily translate abstract ideas into concrete formats, which may not align with natural cognitive processes of creative ideation and refinement [6, 7, 23]. As a result, workflows can become rigid, limiting iterative discovery in generative tasks.

To address this, we introduce **THEMATICPLANE**, a system that supports interactions within thematic design planes, which scaffolds the representation of multidimensional space for exploring and lets users manipulate thematic attributes of visual content. It enables users to ‘ground’ and navigate high-level semantic concepts, such as mood, style, narrative, or emotional tone, in an exploratory fashion. This plane serves as a scaffolding layer, connecting users’ tacit intent with system capabilities with a ‘plane of existence’ for themes, drawing on semantic interaction paradigms to support creativity without externalization barriers.



**Figure 2: Visualization of tacit/thematic and low-level/technical edits in images.**

## 1.1 System Prototype

THEMATICPLANE implements the thematic design plane as an interactive framework adapted from the Circumplex Model [19] for semantic image editing, enabling users to explore and manipulate high-level thematic attributes without explicit low-level inputs. Given an input image by the user, GPT-4o extracts keywords, which are then refined by removing object-based descriptors and retaining only thematic elements like mood or style. Next, we generate 12 thematic perturbations for each theme, assigning them to left or right directions along a semantic axis. Embeddings are computed using DINOv2 [18] for the input image and its aligned text encoder [15] for textual themes, ranked by cosine similarity. Ranked descriptors are then injected into prompts for the image generation model (*i.e.*, Imagen 3 [3]) to produce transformed images. After each regeneration, users can choose to make the new image the primary reference to continue editing.

## 2 Exploratory Study

We recruited 6 participants (4 women, 2 men), aged 19–31, through a corporate Slack channel. Participants had prior experience with image generation tools (e.g., Midjourney [13]). They were first provided 10 minutes to explore the system, using their own images or study provided examples. We then designed three image transformation tasks requiring participants to modify source images to match target images. Each participant completed tasks across three conditions: (1) baseline ChatGPT constrained to image generation only, (2) ChatGPT with full features, and (3) our system—with 5 minutes per condition, followed by a semi-structured interview and survey. Each participant received \$20 for their participation.

## 3 Discussion and Future Work

**Aligning Interfaces to the Human Creativity Process.** Our study revealed two emergent behaviors—open-ended discovery (*i.e.*, divergent exploration) and goal-based editing (*i.e.*, convergent refinement) [10, 25]. Although participants using THEMATICPLANE did not always anticipate the resulting images ( $M = 3.43$ ,  $SD = 1.27$ ), they reported high satisfaction with the outcomes ( $M = 5.86$ ,  $SD = 0.90$ ). Participants noted that non-linear generations either spurred intentional shifts in design direction during divergent tasks (P3) or enabled iteration in a space where themes amalgamated into cohesive affordances not characterized by a single plane (P6).

**Different Perceptions of Thematic Planes.** We observed that participants grounded themselves in themes based on their own understanding, enabling them to interpolate and navigate the rest of the thematic design plane; however, such understanding varied across participants, leading to differing expectations during exploration. Specifically, some participants (P1, P3, P7) anticipated linear mappings between themes and outputs, while others (P2, P4, P5) encountered unclear connections between the plane and initial themes, prompting a desire for explainable methods to comprehend steering mechanisms, rather than merely executing them [2, 12].

**Interactive Visualizations of Themes.** Despite the steerability it offers, our interaction method currently provides a single axis for users to steer generative outputs based on an initial thematic attribute, which may restrict the exploration of interconnected

semantic dimensions [12]. Future work could advance this to multidimensional representations, such as interactive planes or graphs, enabling users to simultaneously manipulate multiple attributes (e.g., mood and style) through gestures like dragging or clustering for more nuanced control over latent space trajectories [22].

**Steerable Model Generations for Images.** Building on the visual representation of steerable dimensions, these should reflect the technical methods for navigating these spaces, extending text-based approaches like LM-Steer [11] to incorporate vector-based steering in multimodal generative systems [8, 14] to create hybrid navigation mechanisms that dynamically adjust trajectories across semantic attributes during inference.

## References

- [1] Krzysztof Adamkiewicz, Paweł W. Woźniak, Julia Dominiak, Andrzej Romanowski, Jakob Karolus, and Stanislav Frolov. 2025. PromptMap: An Alternative Interaction Style for AI-Based Image Generation. In *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*. Association for Computing Machinery, New York, NY, USA, 1162–1176. doi:10.1145/3708359.3712150
- [2] Goomreet Bajaj, Valerie L. Shalin, Srinivasan Parthasarathy, and Amit Sheth. 2024. Grounding From an AI and Cognitive Science Lens. *IEEE Intelligent Systems* 39, 2 (March 2024), 66–71. doi:10.1109/mis.2024.3366669
- [3] Jason Baldridge, Jakob Bauer, Mukul Bhutani, Nicole Brichtova, Andrew Bunner, Lluis Castrejon, Kelvin Chan, Yichang Chen, Sander Dieleman, Yuqing Du, Zach Eaton-Rosen, Hongliang Fei, Nando de Freitas, Yilin Gao, Evgeny Gladchenko, Sergio Gómez Colmenarejo, Mandy Guo, Alex Haig, Will Hawkins, Hexiang Hu, Huilian Huang, Tobienna Peter Igwe, Christos Kaplanis, Siavash Khodadadeh, Yelin Kim, Ksenia Konyushkova, Karol Langner, Eric Lau, Rory Lawton, Shixin Luo, Soňa Mokrá, Henna Nandwani, Yasumasa Onoe, Aáron van den Oord, Zarana Parekh, Jordi Pont-Tuset, Hang Qi, Rui Qian, Deepak Ramachandran, Poorva Rane, Abdullah Rashwan, Ali Razavi, Robert Riachi, Hansa Srinivasan, Srivatsan Srinivasan, Robin Strudel, Benigno Uria, Oliver Wang, Su Wang, Austin Waters, Chris Wolff, Auriel Wright, Zhisheng Xiao, Hao Xiong, Keyang Xu, Marc van Zee, Junlin Zhang, Katie Zhang, Wenlei Zhou, Konrad Zolna, Olá Abuobakar, Canfer Akbulut, Oscar Akerlund, Isabela Albuquerque, Nina Anderson, Marco Andreetto, Lora Aroyo, Ben Bariach, David Barker, Sherry Ben, Dana Berman, Courtney Biles, Irina Blok, Pankil Botadra, Jenny Brennan, Karla Brown, John Buckley, Rudy Bunel, Elie Burszttein, Christina Butterfield, Ben Caine, Viral Carpenter, Norman Casagrande, Ming-Wei Chang, Solomon Chang, Shamik Chaudhuri, Tony Chen, John Choi, Dmitry Churbanau, Nathan Clement, Matan Cohen, Forrester Cole, Mikhail Dektiarev, Vincent Du, Praneet Dutta, Tom Eccles, Ndidi Elue, Ashley Feden, Shlomi Fruchter, Frankie Fricker, Roopal Garg, Weina Ge, Ahmed Ghazy, Bryant Gipson, Andrew Goodman, Dawid Górný, Sven Gowal, Khyatti Gupta, Yoni Halpern, Yena Han, Susan Hao, Jamie Hayes, Jonathan Heek, Amir Hertz, Ed Hirst, Emiel Hoogeboom, Tingbo Hou, Heidi Howard, Mohamed Ibrahim, Dirichi Ike-Njoku, Joana Iljazi, Vlad Ionescu, William Isaac, Reena Jana, Gemma Jennings, Donovan Jenson, Xuhui Jia, Kerry Jones, Xiaoen Ju, Ivana Kajic, Christos Kaplanis, Burcu Karagol Ayan, Jacob Kelly, Suraj Kothawade, Christina Kouridi, Ira Ktena, Jolanda Kumakaw, Dana Kurniawan, Dmitry Lagun, Lily Lavitas, Jason Lee, Tao Li, Marco Liang, Maggie Li-Calisi, Yuchi Liu, Javier Lopez Alberca, Matthieu Kim Lorrain, Peggy Lu, Kristian Lum, Yukun Ma, Chase Malik, John Mellor, Thomas Mensink, Inbar Mosseri, Tom Murray, Aida Nematzadeh, Paul Nicholas, Signe Nørly, Joáo Gabriel Oliveira, Guillermo Ortiz-Jimenez, Michaela Paganini, Tom Le Paine, Roni Paiss, Alisia Parrish, Anne Peckham, Vikas Peswani, Igor Petrovski, Tobias Pfaff, Alex Pirozhenko, Ryan Poplin, Utsav Prabhu, Yuan Qi, Matthew Rahtz, Cyrus Rashtchian, Charvi Rastogi, Amit Raul, Ali Razavi, Sylvestre-Alvise Rebuffi, Susanna Ricco, Felix Riedel, Dirk Robinson, Pankaj Rohatgi, Bill Rosgen, Sarah Rumbley, Moonkyung Ryu, Anthony Salgado, Tim Salimans, Sahil Singla, Florian Schroff, Candice Schumann, Tanmay Shah, Eleni Shaw, Gregory Shaw, Brendan Shillingford, Kaushik Shivakumar, Dennis Shtatnov, Zach Singer, Evgeny Sluzhaev, Valerii Sokolov, Thibault Sottiaux, Florian Stemberg, Brad Stone, David Stutz, Yu-Chuan Su, Eric Tabellion, Shuai Tang, David Tao, Kurt Thomas, Gregory Thornton, Andeep Toor, Cristian Udrescu, Aayush Upadhyay, Cristina Vasconcelos, Alex Vasiloff, Andrey Voynov, Amanda Walker, Luyu Wang, Miaosen Wang, Simon Wang, Stanley Wang, Qifei Wang, Yuxiao Wang, Agoston Weisz, Olivia Wiles, Chenxia Wu, Xingyu Federico Xu, Andrew Xue, Jianbo Yang, Luo Yu, Mete Yurtoglu, Ali Zand, Han Zhang, Jiageng Zhang, Catherine Zhao, Adilet Zhaxybay, Miao Zhou, Shengqi Zhu, Zhenkai Zhu, Dawn Bloxwich, Mahyar Bordbar, Luis C. Cobo, Eli Collins, Shengyang Dai, Tulsee Doshi, Anca Dragan, Douglas Eck, Demis Hassabis, Sissie Hsiao, Tom Hume, Koray Kavukcuoglu, Helen King, Jack Krawczyk, Yeqing Li, Kathy Meier-Hellstern, Andras Orban, Yury Pinsky, Amar Subramanya, Oriol Vinyals, Ting Yu, and Yori Zwols. 2024. Imagen 3. arXiv:2408.07009 [cs.CV] <https://arxiv.org/abs/2408.07009>
- [4] Yali Bian and Chris North. 2021. DeepSI: Interactive Deep Learning for Semantic Interaction. In *Proceedings of the 26th International Conference on Intelligent User Interfaces (College Station, TX, USA) (IUI '21)*. Association for Computing Machinery, New York, NY, USA, 197–207. doi:10.1145/3397481.3450670
- [5] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. arXiv:2312.11949 [cs.HC] <https://arxiv.org/abs/2312.11949>
- [6] DaEun Choi, Kihoon Son, HyunJoon Jung, and Juho Kim. 2025. Expanda: Broadening Design Exploration with Text-to-Image Model. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 232, 10 pages. doi:10.1145/3706599.3720189
- [7] Alex Endert. 2014. Semantic Interaction for Visual Analytics: Toward Coupling Cognition and Computation. *Computer Graphics and Applications, IEEE* 34 (07 2014), 8–15. doi:10.1109/MCG.2014.73
- [8] Patrick Esser, Robin Rombach, and Björn Ommer. 2021. Taming Transformers for High-Resolution Image Synthesis. arXiv:2012.09841 [cs.CV] <https://arxiv.org/abs/2012.09841>
- [9] Rohit Gandikota, Joanna Materzynska, Tingrui Zhou, Antonio Torralba, and David Bau. 2023. Concept Sliders: LoRA Adaptors for Precise Control in Diffusion Models. arXiv:2311.12092 [cs.CV] <https://arxiv.org/abs/2311.12092>
- [10] J.P. Guilford. 1959. Three faces of intellect. *American Psychologist* 14, 8 (1959), 469–479. doi:10.1037/h0046827
- [11] Chi Han, Jialiang Xu, Manling Li, Yi Fung, Chenkai Sun, Nan Jiang, Tarek Abdelzaher, and Heng Ji. 2024. Word Embeddings Are Steers for Language Models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Lun-Wei Ku, Andre Martins, and Vivek Srikumar (Eds.). Association for Computational Linguistics, Bangkok, Thailand, 16410–16430. doi:10.18653/v1/2024.acl-long.864
- [12] Yuanning Han, Ziyi Qiu, Jiale Cheng, and RAY LC. 2024. When Teams Embrace AI: Human Collaboration Strategies in Generative Prompting in a Creative Design Task. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 176, 14 pages. doi:10.1145/3613904.3642133
- [13] Midjourney Inc. 2025. Midjourney. <https://www.midjourney.com/>
- [14] Ali Jahanian, Lucy Chai, and Phillip Isola. 2020. On the “steerability” of generative adversarial networks. arXiv:1907.07171 [cs.CV] <https://arxiv.org/abs/1907.07171>
- [15] Cijo Jose, Théo Moutakanni, Dahuyn Kang, Federico Baldassarre, Timothée Darcret, Hu Xu, Daniel Li, Marc Szafraniec, Michaël Ramamjisoa, Maxime Oquab, et al. 2025. Dinov2 meets text: A unified framework for image-and pixel-level vision-language alignment. In *Proceedings of the Computer Vision and Pattern Recognition Conference*. 24905–24916.
- [16] Kevin Meng, David Bau, Alex Andonian, and Yonatan Belinkov. 2023. Locating and Editing Factual Associations in GPT. arXiv:2202.05262 [cs.CL] <https://arxiv.org/abs/2202.05262>
- [17] Jonas Oppenlaender, Rhema Linder, and Johanna Silvennoinen. 2024. Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering. arXiv:2303.13534 [cs.HC] <https://arxiv.org/abs/2303.13534>
- [18] Maxime Oquab, Timothée Darcret, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Hazziza, Francisco Massa, Alaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jegou, Julien Maijal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. 2024. DINov2: Learning Robust Visual Features without Supervision. arXiv:2304.07193 [cs.CV] <https://arxiv.org/abs/2304.07193>
- [19] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.
- [20] Jingyu Shi, Rahul Jain, Hyungjun Doh, Ryo Suzuki, and Karthik Ramani. 2024. An HCI-Centric Survey and Taxonomy of Human-Generative-AI Interactions. arXiv:2310.07127 [cs.HC] <https://arxiv.org/abs/2310.07127>
- [21] Kihoon Son, DaEun Choi, Tae Soo Kim, Young-Ho Kim, and Juho Kim. 2024. GenQuery: Supporting Expressive Visual Search with Generative Models. arXiv:2310.01287 [cs.HC] <https://arxiv.org/abs/2310.01287>
- [22] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 644, 26 pages. doi:10.1145/3613904.3642400
- [23] Michael Terry, Elizabeth D. Mynatt, Kumiko Nakakoji, and Yasuhiro Yamamoto. 2004. Variation in element and actions supporting simultaneous development of alternative solutions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Vienna, Austria) (CHI '04)*. Association for Computing Machinery, New York, NY, USA, 711–718. doi:10.1145/985692.985782

- [24] Zhijie Wang, Yuheng Huang, Da Song, Lei Ma, and Tianyi Zhang. 2024. PromptCharm: Text-to-Image Generation through Multi-modal Prompting and Refinement. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 185, 21 pages. doi:10.1145/3613904.3642803
- [25] Robin H. Willemse, Isabelle C. de Vink, Evelyn H. Kroesbergen, and Ard W. Lazonder. 2023. The role of creative thinking in children's scientific reasoning. *Thinking Skills and Creativity* 49 (2023), 101375. doi:10.1016/j.tsc.2023.101375