Cognitive Load and Creative Flow in Dynamic Poster Design: A Comparative Study of Manual, AI-Generated, and Human-AI Collaborative Workflows

Lin Zhang, Weiqian Mao*, Yingfei Jia, Yang Liu

Harbin Institute of Information Technology, Harbin, China

*642559537@qq.com

Abstract. Dynamic posters require coordinated visual hierarchy, motion timing, and rapid iteration. Generative AI can support drafts, style references, and motion concepts, but pure AI workflows often shift effort to prompt refinement, selection, and correction. Rather than proposing a new model, this paper presents a reproducible framework for comparing manual, pure-AI, and human-AI workflows in dynamic poster design. The framework standardizes task conditions, AI tools, participant expertise, and a normalized effectiveness index combining cognitive load, creative flow, quality, and completion time. A designed experiment with 60 participants is outlined. Expected results suggest that human-AI collaboration reduces unnecessary load while preserving designer control and creative engagement.

Keywords: Dynamic poster design; Generative AI; Human-AI collaboration; Cognitive load; Creative flow; Multimedia design.

1 Introduction

Dynamic posters combine typography, imagery, and motion in a compact format, requiring designers to translate briefs into visual hierarchy, rhythm, and brand expression under time constraints. Vision-language and diffusion models enable rapid exploration, but effective design still depends on intention control, editability, and motion organization. Pure AI workflows offer speed but can be unstable, while manual design provides control at higher cost. This paper proposes a workflow-level evaluation framework, comparing manual, pure-AI, and human-AI workflows in terms of cognitive load, creative flow, output quality, and efficiency..

2 Literature Review

2.1 Generative AI in Visual and Dynamic Design

Generative AI has changed how visual concepts are produced. CLIP-style models align text and image representations, while DALL-E-style and diffusion systems

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_41

transform prompts into coherent images with semantic richness and stylistic variety [1][2][3][4]. In design practice, these systems support mood boards, references, color schemes, and alternative compositions before motion production begins.

However, current generative models also introduce workflow problems. Outputs remain probabilistic and often require prompt adjustment, selection, retouching, and format conversion. For time-based poster design, a single image does not provide editable layers, motion logic, brand rules, or transition timing. AI examples may narrow idea diversity if users follow them too closely [5], and recent work shows that design experience and interface strategy affect creativity and cognitive load .

2.2 Human-AI Collaboration, Cognitive Load, and Creative Flow

Human-AI collaboration research suggests that productive creative work requires more than one-directional instruction. Co-design frameworks emphasize exploration, reflection, and remixing, where AI expands possibilities rather than only executing commands [6]. Studies of co-drawing, design expertise, and agency-first interaction show that AI cooperation changes attention, perceived control, self-efficacy, and cognitive load .

Cognitive load and creative flow are therefore useful indicators for evaluating AI-assisted design. Cognitive load reflects task understanding, tool operation, selection, and repair effort; creative flow reflects concentration, control, feedback, and enjoyment. Existing studies still lack a compact, replicable framework for comparing manual, pure-AI, and human-AI workflows under the same dynamic poster task.

3 Methodology

3.1 Overall Comparative Framework

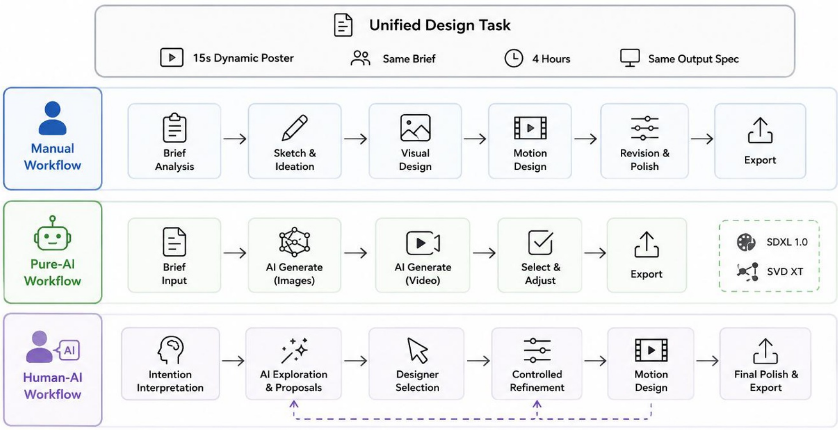

Fig. 1. Workflow comparison for dynamic poster design.

As shown in Fig. 1, the proposed framework compares three workflow modes under a unified dynamic poster task. The manual workflow covers brief analysis, sketching, production, animation, and revision without generative AI. The pure-AI workflow uses AI as the primary source of visual and motion concepts, with only necessary selection and export adjustment. The human-AI workflow organizes the process into intention interpretation and iterative refinement, AI-supported exploration, designer selection, controlled refinement, and manual motion organization. Here, intention interpretation is understood as an iterative cognitive process in which designers continuously interpret and adjust design goals in response to AI-generated outputs. The framework treats the diffusion model as a fixed component and provides a reproducible role-allocation and evaluation protocol.

3.2 Workflow Definition and Task Control

The task is to design a 15-second dynamic poster for a fictional brand. This duration is selected as a controlled and representative time span commonly used in short-form motion design, allowing key elements such as visual hierarchy, transitions, and rhythm to be meaningfully expressed within a manageable scope. It balances experimental control and ecological validity by ensuring sufficient design complexity while keeping task demands comparable across participants. While different durations may influence cognitive load (e.g., longer tasks increasing planning demands), the duration is fixed here to isolate workflow effects; future work may examine duration variations.

Each participant receives the same brief, including audience, message, required text, brand tone, resolution, and delivery format. The time limit is four hours. Manual participants use non-generative image-editing and animation software. Pure-AI participants use ComfyUI with SDXL 1.0 for key visuals and Stable Video Diffusion XT for motion proposals, performing only selection, trimming, and export adjustment. Human-AI participants use the same AI tools for exploration, while final composition, hierarchy, rhythm, and motion remain designer-controlled.

To ensure comparability, all groups use the same asset package, font set, export standard, and evaluation rubric. For replication, AI-based workflows record key parameters including model checkpoint, ComfyUI workflow, seed, prompts, sampling steps, resolution, and post-editing actions. Screen recordings and logs capture prompt iterations, editing operations, selection time, and revision loops..

3.3 Evaluation Variables and Quantitative Model

Four primary indicators are adopted. Cognitive load is measured by NASA-TLX after task completion. Creative flow is measured through a short scale for concentration, control, feedback, and enjoyment. Output quality is scored by three experts using a five-point rubric for hierarchy, originality, motion coherence, brand consistency, and technical completion. Efficiency is measured by completion time. Raw measures are min-max normalized: Q_n and F_n increase with higher quality and flow, while L_n

and T_n increase with higher load and longer time. The normalized effectiveness index is:

$$W_eff = \alpha Q_n + \beta F_n + \gamma(1 - L_n) + \delta(1 - T_n) \tag{1}$$

where Q_n is normalized expert quality, F_n is normalized creative flow, L_n is normalized cognitive load, and T_n is normalized completion time. Because lower load and shorter time are better, L_n and T_n enter as 1-L_n and 1-T_n. Equal weights are used here to avoid overemphasizing speed or visual quality alone.

A weighted cognitive-load estimate is also used to describe the source of effort during the process:

$$L_total = w_i L_intent + w_o L_operation + w_s L_selection + w_r L_repair \tag{2}$$

where L_{intent} refers to understanding and translating the design intention, $L_{\text{operation}}$ to tool operation, $L_{\text{selection}}$ to choosing among alternatives, and L_{repair} to correcting unusable outputs. This model is important because pure-AI workflows may reduce operation effort while increasing selection and repair load, whereas human-AI collaboration is expected to reduce L_{repair} by keeping the designer in control of structure and final decisions.

This decomposition aligns with classical cognitive load theory: intention understanding reflects intrinsic load; selection and correction correspond to extraneous load from tool and output limitations; and tool operation includes both extraneous and germane load when it supports schema construction and design refinement, especially in human-AI workflows.

3.4 Experimental Procedure and Data Processing

The planned experiment recruits 60 participants with basic dynamic design experience, including senior students and junior designers. To control baseline differences, participants complete an expertise pretest covering years of practice (years), animation-software familiarity (Likert 1–5), AI-tool experience (Likert 1–5), and a portfolio rating (expert score 1–5). All components are z-score normalized, and the expertise score E is computed as their average. Stratified random assignment ensures comparable E distributions across three groups (n = 20 each). All participants receive the same tool familiarization and export requirements.

Participants work independently and, upon completion, submit their posters, questionnaires, and brief workflow reflections. Expert scoring is anonymous. Indicators are analyzed using descriptive statistics and ANOVA (when data are collected), with E reported as a covariate or stratified descriptor.

To complement NASA-TLX results, qualitative reflections are analyzed via thematic coding. Responses are segmented and coded based on the four components of the weighted cognitive-load model (intention, operation, selection, correction), with allowance for emergent categories. Two coders perform the analysis, and inter-rater reliability is assessed using Cohen's kappa.

Coded themes are mapped onto NASA-TLX dimensions (e.g., mental demand, effort, frustration) to enable triangulation between ratings and reported experiences, clarifying how cognitive load sources vary across workflows.

As this paper reports a designed experiment, the next section presents expected results based on workflow logic and prior findings..

4 Experiments and Comparative Analysis

4.1 Experimental Setup

The setup is a between-group comparison. The manual group completes layout design, image editing, key-frame animation, and export adjustment. The pure-AI group uses the predefined ComfyUI, SDXL 1.0, and Stable Video Diffusion XT workflow as the primary production method. The human-AI group uses the same AI tools for ideation and variation, but manually controls composition, rhythm, and final polish. All groups complete the same 15-second poster within four hours.

The expected evaluation combines subjective and expert measures. NASA-TLX represents L_n, the creative-flow scale represents F_n, expert quality score represents Q_n, and completion time represents T_n after normalization. This setup reveals whether efficiency improves creativity or shifts effort to prompt repair and output selection.

4.2 Main Expected Experimental Results

Table 1. Expected comparative results of three dynamic poster workflows.

Workflow	NASA-TLX	Creative flow	Quality score	Completion time
Manual	58.4	4.05 / 5	4.12 / 5	228 min
Pure-AI	64.7	3.28 / 5	3.62 / 5	176 min
Human-AI	46.9	4.43 / 5	4.28 / 5	158 min

Table 1 presents the anticipated pattern rather than completed empirical data. The human-AI workflow is expected to achieve the lowest cognitive load and highest creative flow because AI supports exploration while the designer controls intention, hierarchy, and motion. For consistency with Eq. (1), raw values would be converted into Q_n, F_n, L_n, and T_n before W_eff is calculated; larger L_n and T_n reduce W_eff through 1-L_n and 1-T_n. The pure-AI group should be faster than the manual group, but its load may remain high because participants must adjust prompts, screen unstable outputs, and repair inconsistencies.

4.3 Discussion

The expected advantage of human-AI collaboration comes from role allocation rather than from a new generative model. AI expands visual alternatives and accelerates exploration, but it is less reliable when asked to independently complete a controlled communication task. Human-AI collaboration combines AI possibility generation with human judgment for semantic control, motion organization, and quality assurance.

The framework also clarifies why pure-AI production is not necessarily the lowest-load workflow. If generated outputs are difficult to edit, designers spend effort on selection and repair rather than creation, reducing control and flow even when completion time is shorter. AI tools should therefore be integrated with transparent settings, logged prompts, and designer-controlled decisions.

5 Conclusion

This study proposed a reproducible workflow-level framework for evaluating manual, pure-AI, and human-AI workflows in dynamic poster design. By combining cognitive load, creative flow, expert quality, completion time, AI-tool logging, participant expertise control, and a consistent normalized effectiveness index, it expands assessment beyond visual output and efficiency alone.

The designed experiment indicates that human-AI collaboration can reduce repair burden, support wider exploration, and preserve control over communication intention and motion structure. Pure-AI generation may improve speed but increase selection and correction load.

Future work will conduct the experiment with real participants, compare AI interfaces, report expertise-stratified results, and extend the framework to interactive posters and short-video advertising.

References

1. Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical Text-Conditional Image Generation with CLIP Latents[EB/OL]. arXiv:2204.06125, 2022. DOI: 10.48550/arXiv.2204.06125.
2. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022:10684-10695. DOI: 10.1109/CVPR52688.2022.01042.
3. Wadinambiarachchi S, Kelly R M, Pareek S, Zhou Q, Velloso E. The Effects of Generative AI on Design Fixation and Divergent Thinking[C]//Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: ACM, 2024. DOI: 10.1145/3613904.3642919.
4. Zhou J, Li R, Tang J, Tang T, Li H, Cui W, Wu Y. Understanding Nonlinear Collaboration between Human and AI Agents: A Co-design Framework for Creative

Design[C]//Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: ACM, 2024. DOI: 10.1145/3613904.3642812.
5. Ho J, Jain A, Abbeel P. Denoising Diffusion Probabilistic Models[C]//Advances in Neural Information Processing Systems (NeurIPS). 2020. DOI: 10.48550/arXiv.2006.11239.
6. Nichol A Q, Dhariwal P. Improved Denoising Diffusion Probabilistic Models[C]//Proceedings of the 38th International Conference on Machine Learning (ICML). 2021. DOI: 10.48550/arXiv.2102.09672.