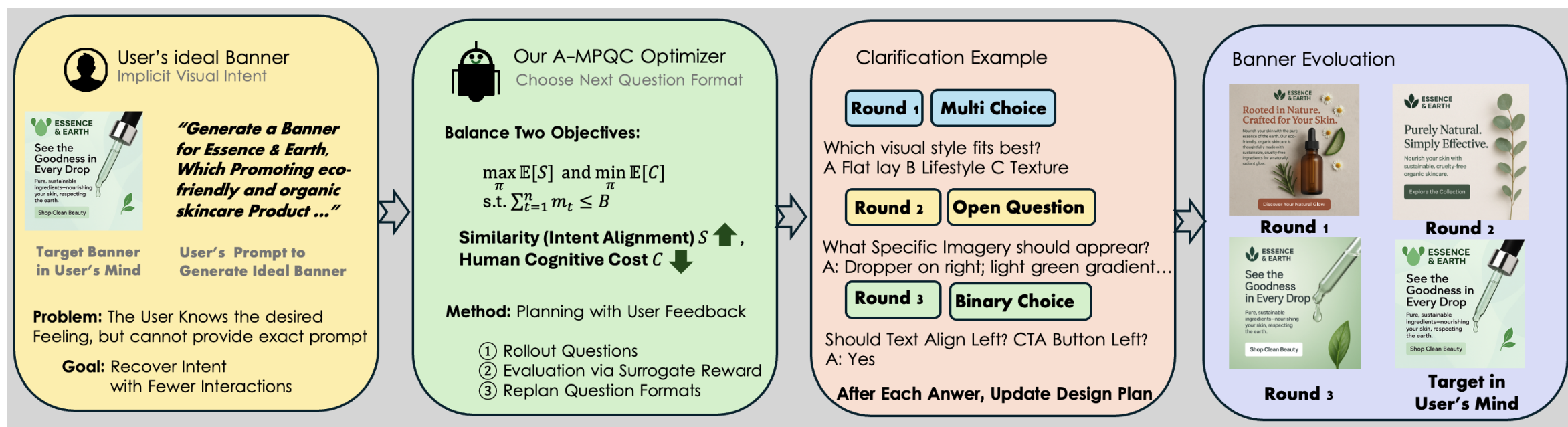


Agentic Model Predictive Questioning Control in Visual Design

**Kuang-Da Wang * , Zhao Wang* (Corresponding Author, Presentor),
Wei-Yao Wang , Yotaro Shimose , Jaechang Kim, Shingo Takamatsu**

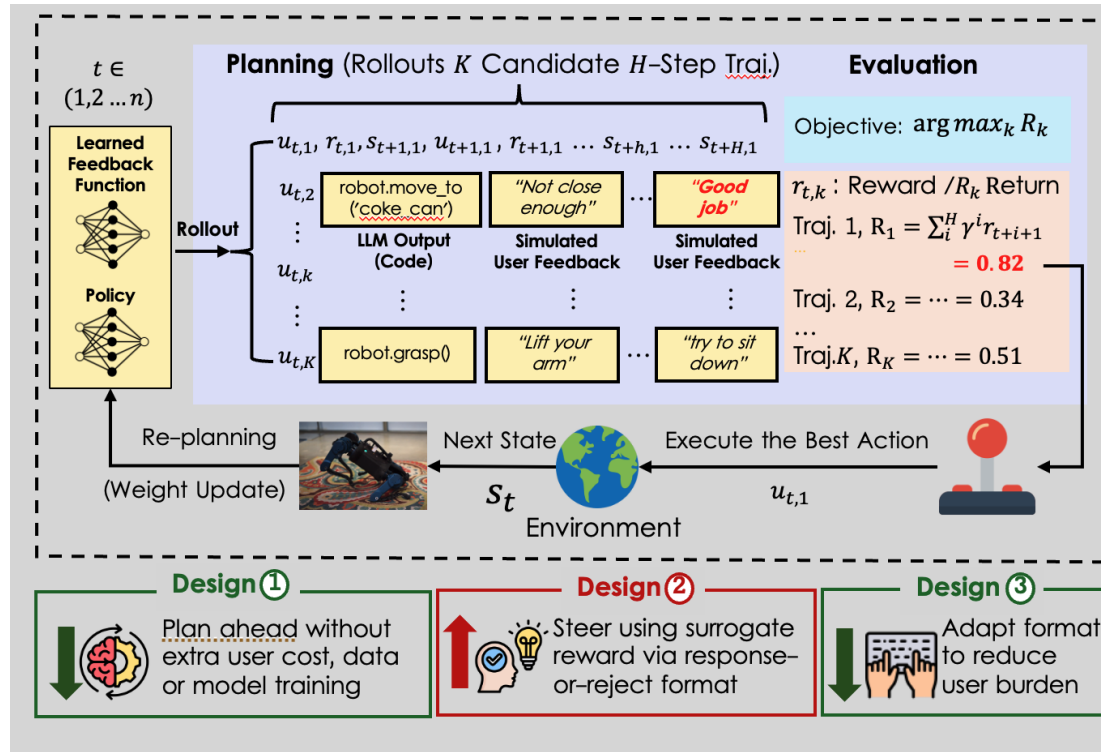
Motivation: Ask Less, Align Better !

Balancing Design Alignment and Human Cognitive Cost in Human-Agent Visual Co-creation Tasks.

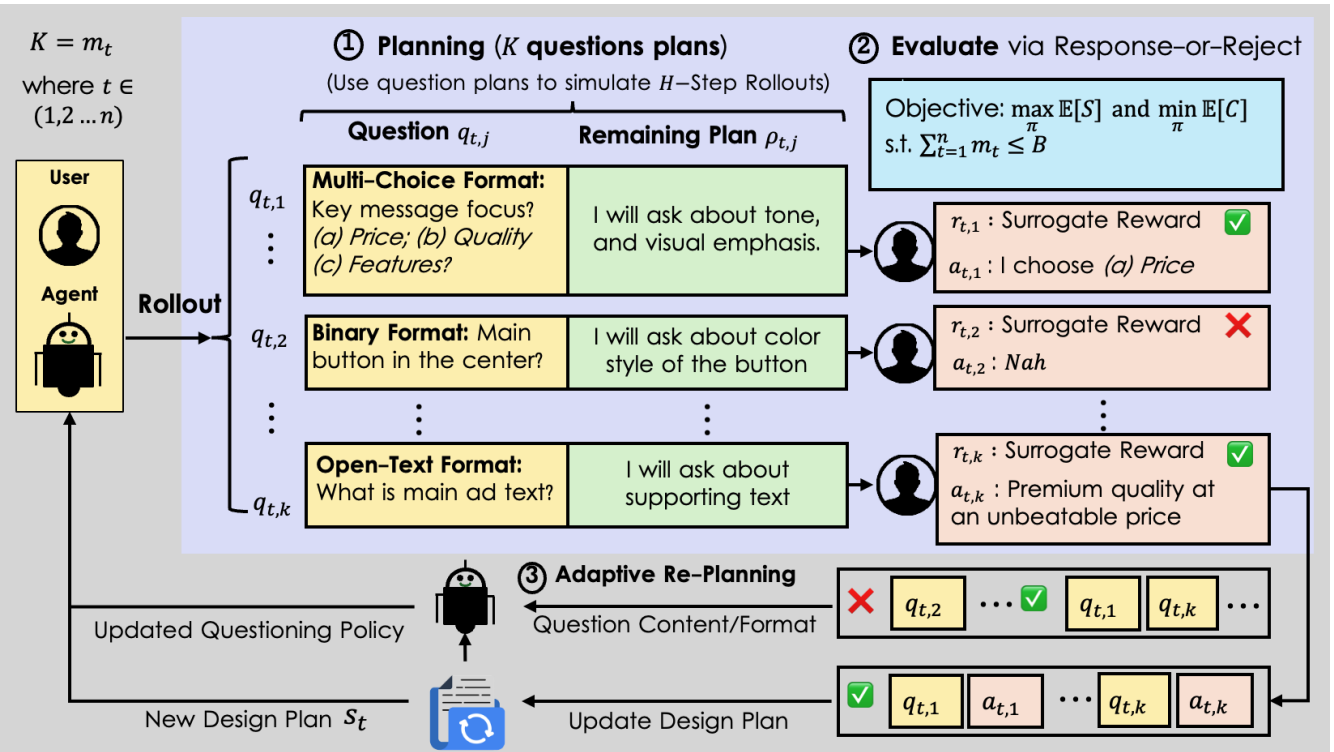


Method: the First Optimization-based Multi-round Clarification for Human-Agent Visual Co-creation

Standard Model Predictive Control with User Feedback



Our Proposal: Model Predictive Questioning Control



Experimental Result

WebGen-V (Web HTML Generation)				MIMO (Banner Image Generation)			
Method	$\tilde{C} \downarrow$	$\Delta S \uparrow$	$\Delta S / \tilde{C} \uparrow$	Method	$\tilde{C} \downarrow$	$\Delta S \uparrow$	$\Delta S / \tilde{C} \uparrow$
No user interaction	0.000	0.000	-	No user interaction	0.000	0.000	-
DG (Directly Generation)	6.046	0.366	0.060	DG (Directly Generation)	6.646	0.256	0.038
DG + Binary	0.116	0.017	<u>0.150</u>	DG + Binary	0.055	0.019	<u>0.342</u>
DG + Multiple-Choice	6.449	0.295	<u>0.045</u>	DG + Multiple-Choice	1.614	0.070	<u>0.043</u>
DG + Open-Ended	13.276	0.566	0.042	DG + Open-Ended	5.240	0.209	0.040
DG + Flexible	9.603	0.418	0.043	DG + Flexible	2.318	<u>0.457</u>	0.197
DG + Flexible + RAG	7.298	0.383	0.052	DG + Flexible + RAG	1.331	0.292	0.219
A-MPQC (Ours)	<u>2.548</u>	<u>0.480</u>	0.188	A-MPQC (Ours)	<u>1.238</u>	0.594	0.480

Does A-MPQC yield better tradeoffs?

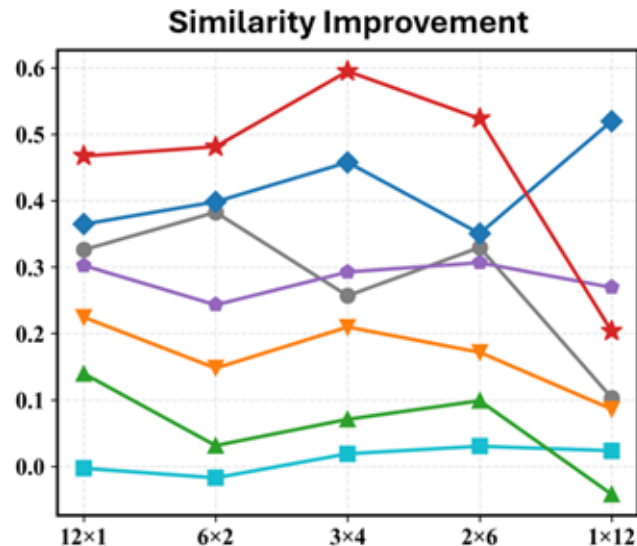


Table 2. Ablation study on MIMO under fixed budget $B=12$ with $(n, m)=(3, 4)$. MPC-on- m is reported as an alternative control dimension (not a design removal).

Variant	$\tilde{C} \downarrow$	$\Delta S \uparrow$	$\Delta S / \tilde{C} \uparrow$
A-MPQC (full)	1.238	0.594	0.480
A-MPQC-on- m	6.696	<u>0.537</u>	0.080
w/o Design I (no question plans)	0.925	0.391	0.422
w/o Design II (all-respond)	0.474	0.204	<u>0.431</u>
w/o Design III (direct questioning)	<u>1.758</u>	0.273	0.155

How should a budget be allocated?

Which components drive A-MPQC's gains?



Essence & Earth. Promoting eco-friendly and organic skincare products that nourish your skin naturally. Highlight the appeal of sustainable and cruelty-free ingredients for a healthier lifestyle.

Before Asking Question
Generate Initial Design Plan



Which visual style best captures the brand's essence for the poster's main image?
(A) A top-down flat lay of the product surrounded by natural ingredients.
(B) A lifestyle scene with a person enjoying the product in a calm, natural setting.
(C) An artistic close-up of a single key ingredient or natural texture.
(D) A simple, clean background with only the product bottle as the focus.

Round 1, Question 1

No Rejected by answerer

Which visual concept best fits the poster?
(A) A product flat lay with ingredients,
(B) A lifestyle photo with a person in nature,
(C) An abstract nature texture.

Round 1, Question 2

No Rejected by answerer

Which typographic style best fits the poster's aesthetic?
(A) Elegant & Classic (Serif headline, clean sans-serif body text)
(B) Modern & Minimalist (All sans-serif fonts)
(C) Natural & Personal (A touch of script font for the headline)
(D) Other (please specify)

Round 1, Question 3

B

Round 1 Finished
Update Design Plan

What specific imagery, background textures, or visual elements should the poster feature to represent your eco-friendly products?

Round 2, Question 1

The main visual element should be a close-up of a glass dropper with a clear serum, positioned on the right side. The background should be a soft, light green gradient, not a texture. Also, include a small, sharp-focus green leaf in the bottom right corner.

Given the imagery will be on the right, should the logo and text be aligned to the left side? (Yes/No)

Round 2, Question 2

Yes

What should the headline, description, and call-to-action text be for the poster?

Round 2, Question 3

The headline is 'See the Goodness in Every Drop', the description is 'Pure, sustainable ingredients—nourishing your skin, respecting the earth.', and the call-to-action is 'Shop Clean Beauty'.

Round 2 Finished
Update Design Plan

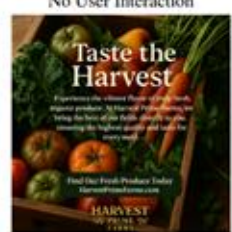
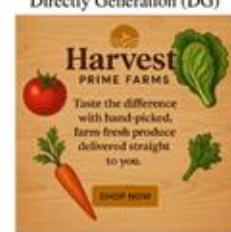
What visual style should the 'Shop Clean Beauty' call-to-action button have?
(A) A solid, filled rectangle for strong emphasis.
(B) A rectangle with rounded corners for a softer, modern look.
(C) A 'ghost' button (an outline with a transparent fill).
(D) Simple underlined text for a very minimalist feel.

Round 3, Question 1

B



Agent Already Satisfy
Generate the Image



Target Image

Ours A-MPQC

Directly Generation (DG)

No User Interaction

Human-Agent Co-Creation Example

SONY

“Sony”はソニー株式会社の商標です。

各ソニー製品の商品名・サービス名はソニー株式会社またはグループ各社の登録商標です。その他の製品および会社名は、各社の商号、登録商標または商標です。