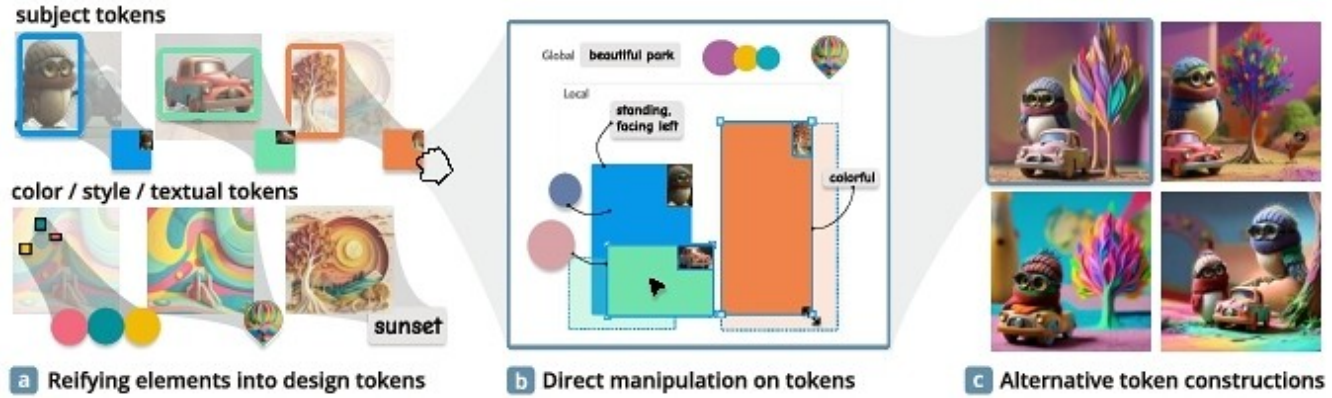
Brickify

Enabling Expressive Design Intent Specification through Direct Manipulation on Design Tokens

CHI 2025 — Xinyu Shi, Yinghou Wang, Ryan Rossi, Jian Zhao

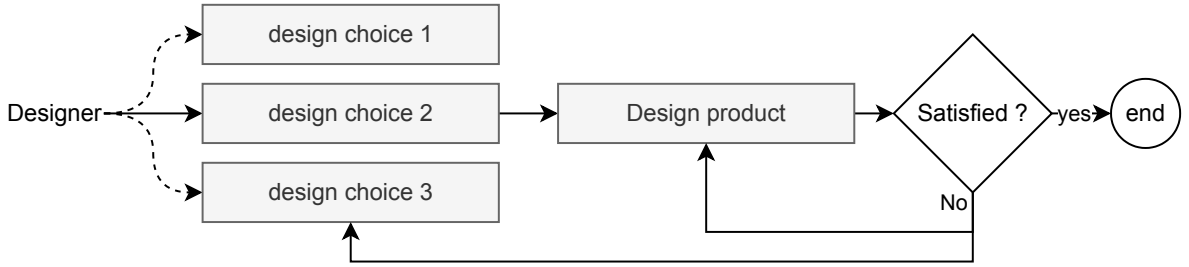
1. INTRODUCTION

Visual abstract

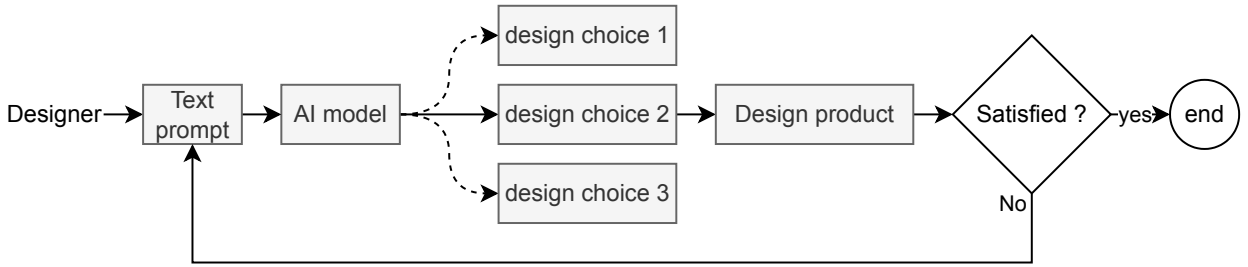


The Design Problem:

Designed process as supposed:



Designed process as proposed by TTI Gen AI:



Is Natural language good at describing visuals?

Is Natural language good at describing visuals?

- Can you name those colors ?

Is Natural language good at describing visuals?

- Can you name those colors ?

Is Natural language good at describing visuals?

- Can you name those colors ?

Is Natural language good at describing visuals?

- Can you name those colors ?

Is Natural language good at describing visuals?

What is the color of the dress?

Is Natural language good at describing visuals?

Is Natural language good at describing visuals?

- Imprecise spatial relations

Is Natural language good at describing visuals?

- Imprecise spatial relations

2. BRICKIFY CONCEPT

Brickify introduces a visual-centric paradigm:

From Text to Tokens

From Text to Tokens

From Text to Tokens

From Tokens to Image

From Tokens to Image

Each visual element becomes a **design token.**

- Extract tokens from reference images

- Manipulate tokens (move, resize, link, group)

- Build a visual lexicon representing intent

Each visual element becomes a **design token.**

- Extract tokens from reference images

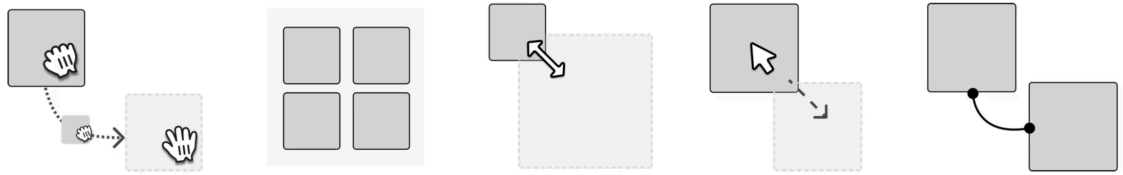
- Manipulate tokens (move, resize, link, group)

- Build a visual lexicon representing intent

Each visual element becomes a **design token.**

- Extract tokens from reference images

- Manipulate tokens (move, resize, link, group)

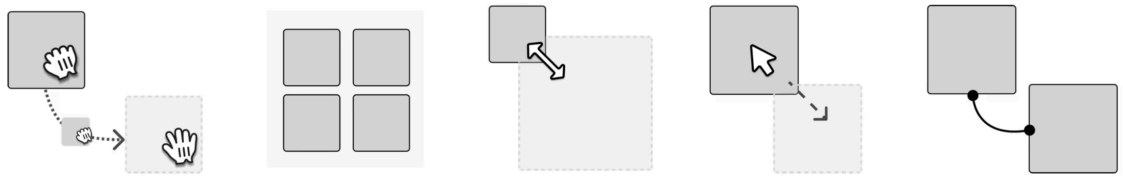


- Build a visual lexicon representing intent

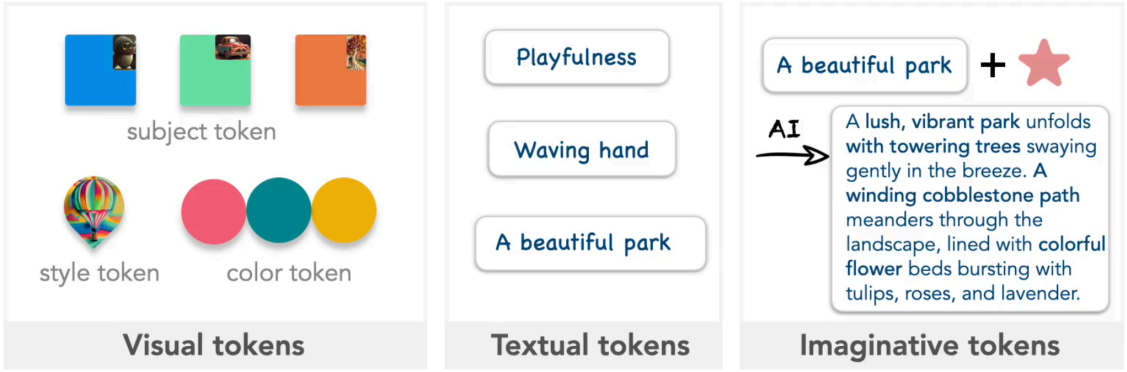
Each visual element becomes a **design token.**

- Extract tokens from reference images

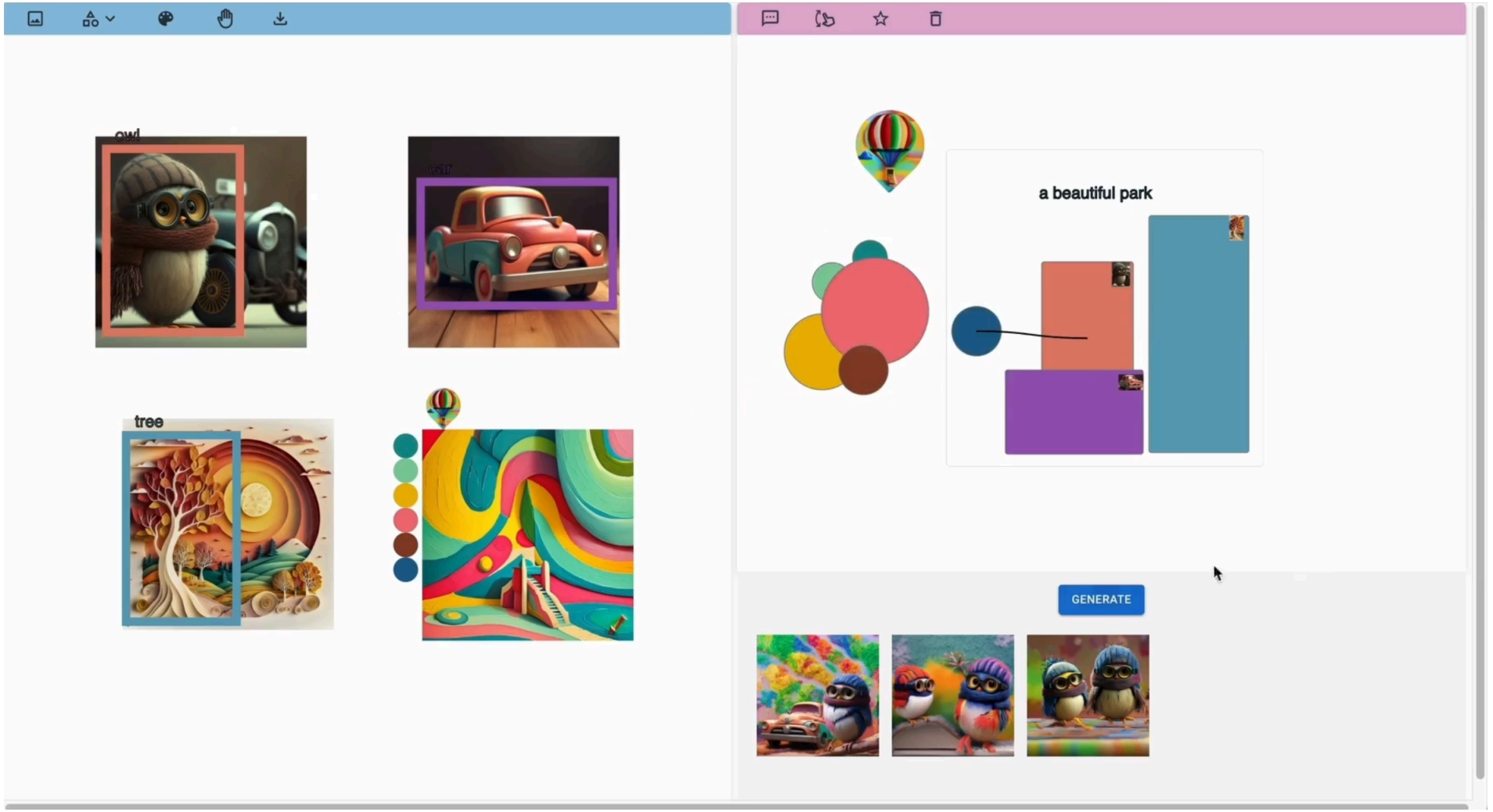
- Manipulate tokens (move, resize, link, group)

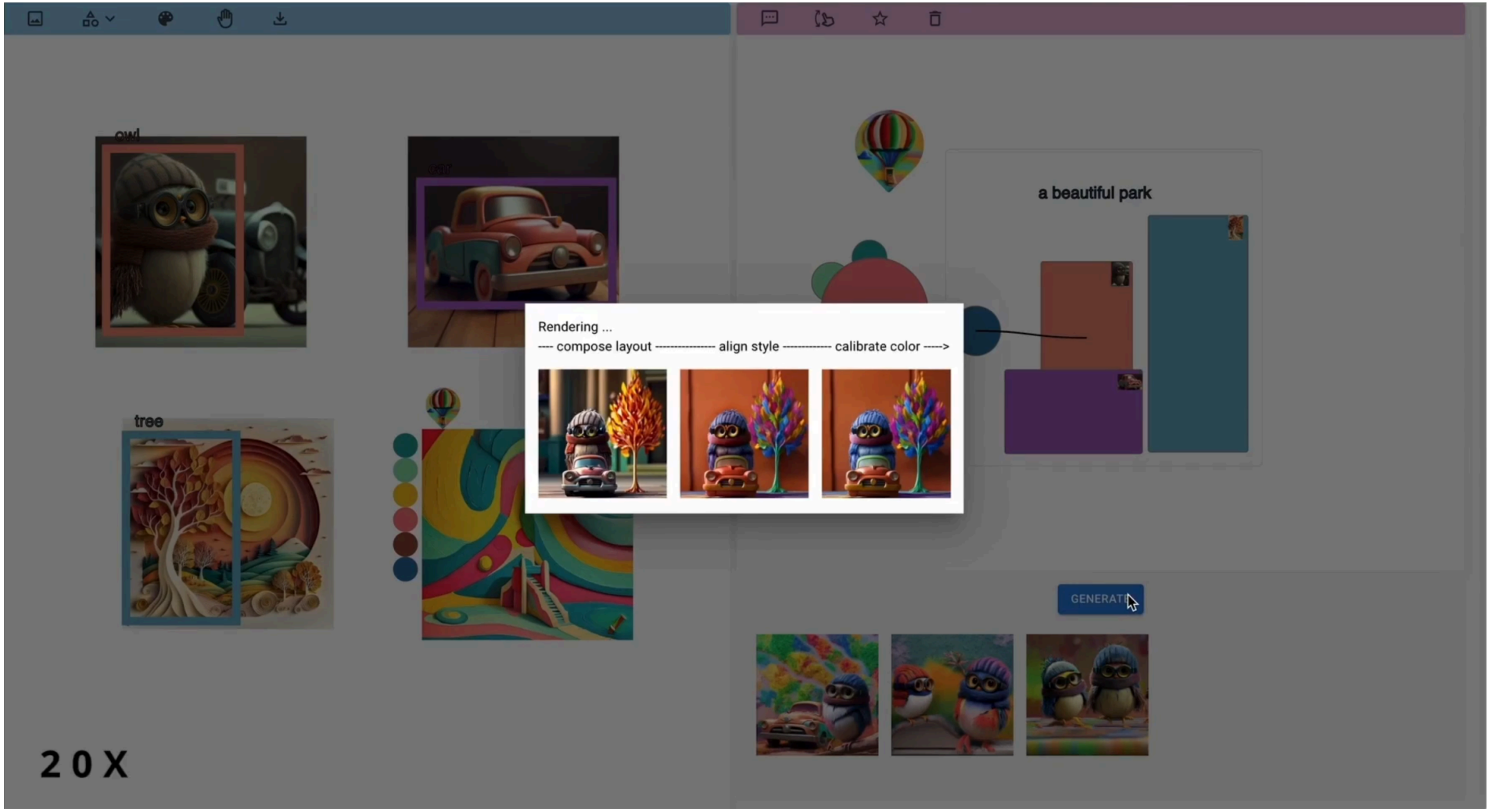


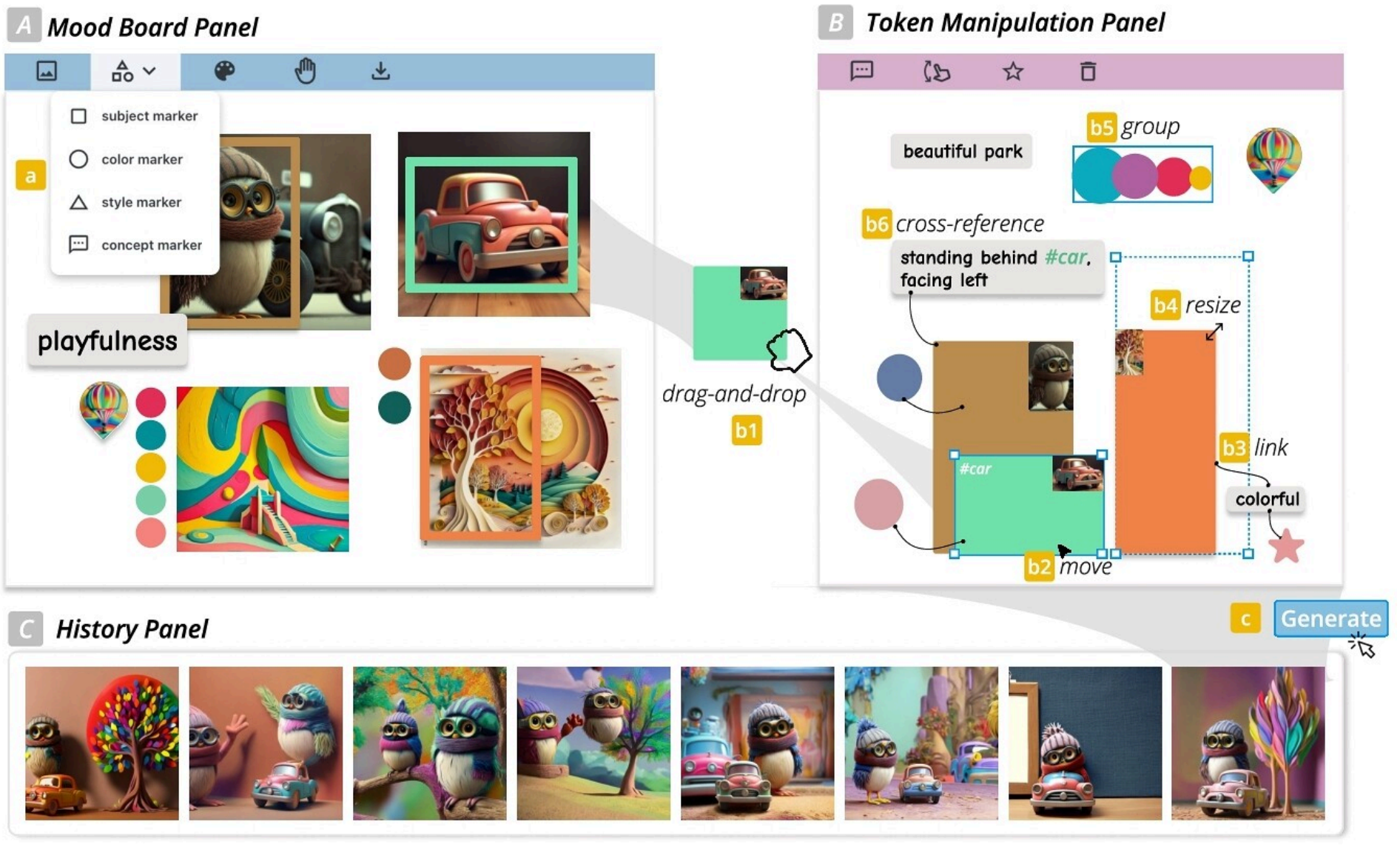
- Build a visual lexicon representing intent



Users control what elements to use — Delegate control to AI


owl
car
tree
a beautiful park
GENERATE


owl
tree
a beautiful park
Rendering ...
---- compose layout ---------- align style ---------- calibrate color ----->
GENERATE
20X


A Mood Board Panel
subject marker
color marker
style marker
concept marker
a
playfulness
drag-and-drop
b1
B Token Manipulation Panel
beautiful park
b5 group
b6 cross-reference
standing behind #car, facing left
b4 resize
b3 link
colorful
#car
b2 move
c
Generate
C History Panel

3. DESIGN PROCESS

Four Stages

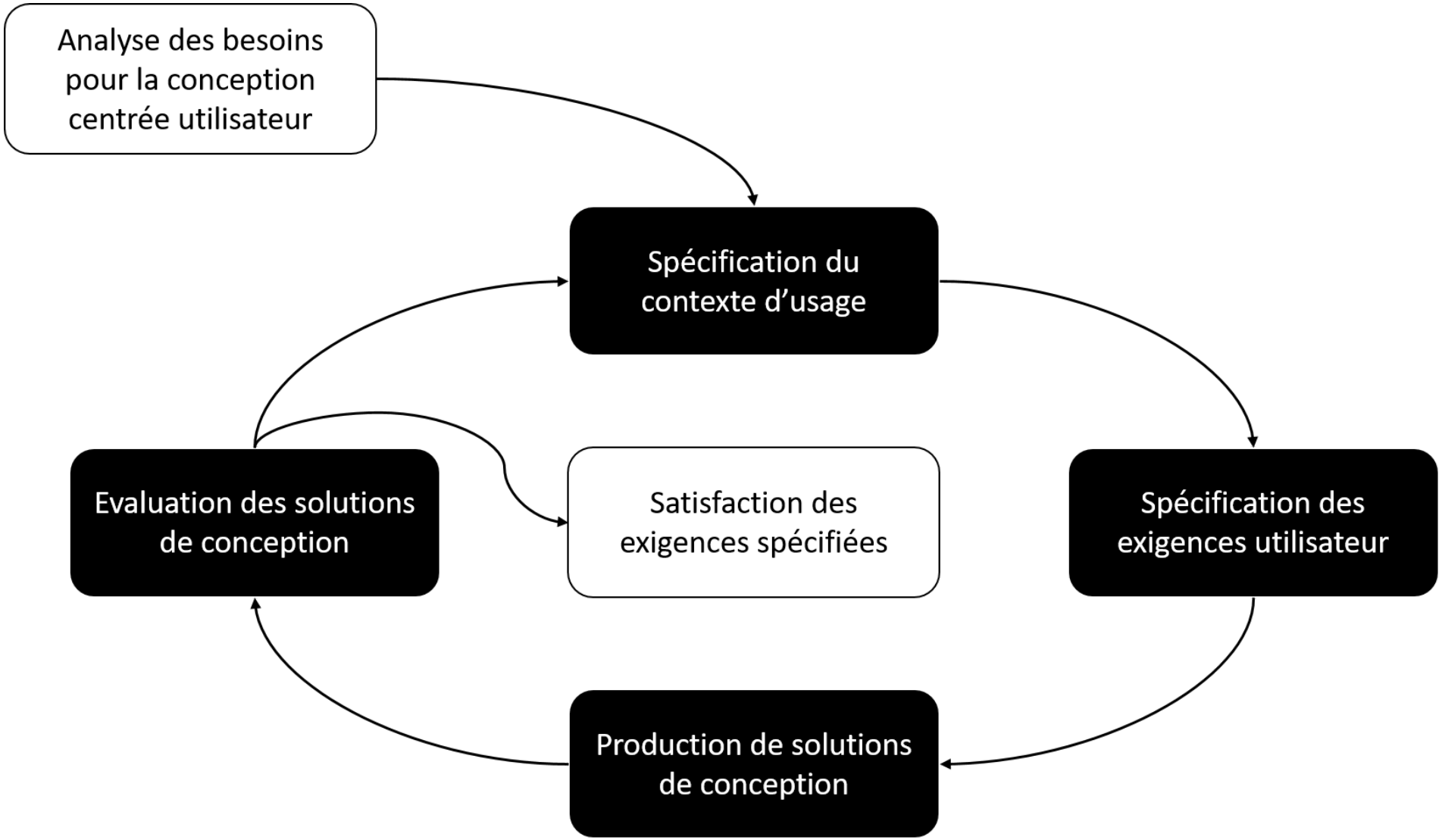
S1 - Problem Understanding: 6 semi-structured interviews with SME

S2 - Early Prototyping: co-design with 1 expert designer

S3 - Iteration: feedback from 6 designers following a user test

S4 - Evaluation: controlled user study (N = 12)

Four Stages of HCD for comparison (IS09241-210)



S1 - Problem Understanding

Identified Challenges and associated Design Goals

S1 - Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI

S1 - Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI
  - DG1: Externalize selective focus

S1 - Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI
  - DG1: Externalize selective focus
- C2: Hard to verbalize relationships

S1 - Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI
  - DG1: Externalize selective focus
- C2: Hard to verbalize relationships
  - DG2: Enable spatial management & visual communication

S1 - Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI
  - DG1: Externalize selective focus
- C2: Hard to verbalize relationships
  - DG2: Enable spatial management & visual communication
- C3: Inefficient iterative refinement

S1 – Problem Understanding

Identified Challenges and associated Design Goals

- C1: Failing to convey attended elements to AI
  - DG1: Externalize selective focus
- C2: Hard to verbalize relationships
  - DG2: Enable spatial management & visual communication
- C3: Inefficient iterative refinement
  - DG3: Facilitate reuse & iteration

S2 - Early Prototyping:

- Weekly 30 minutes design meetings
- Low-fidelity mock-ups
- Non-functional prototype in Figma

S3 - Prototype iteration

User study (n=6) designers

1. Walkthrough of the system
2. Exploratory use of the system with think-aloud verbalizations.

S3 - Prototype iteration

Feedback

1. Strengthen the visual association between design tokens and original imagery to improve clarity
2. Introducing a cross-referencing feature to allow for more effective descriptions of relationships between subject tokens.
3. Added imaginative token to the interaction vocabulary

S4: System evaluation

Controlled user study Participants: (n=12) experienced designers

Two tasks:

1. **Task 1:** replication with a clear intent
2. **Task 2:** open-ended exploration

Task 1 – Comparison

RQ: how does the visual-centric interaction paradigm of Brickify compare to the textual-centric paradigm in terms of clarity, mental effort and time investment for expressing design intent?

2×2 within-subject design:
Technique (Brickify vs Baseline) × **Difficulty** (Easy vs Hard)

STUDY 1 Reference Images

Target Images

Task 1 – Comparison

Metrics:

- human-evaluation (3 external raters using 5 likert scales)
  Element coverage, Size clarity, Position clarity, Style clarity, Color clarity
- Self evaluation of intent expression (5 questions)
- Task time (initial completion and refinement)
- Preferences
- Self reported cognitive load

Task 1 – Results

Participants showed a clear preference for Brickify

Brickify led to:

- ↑ Design intent expression experience
- ↓ Mental effort and frustration
- ↑ Initial completion time
- ↓ Refinement time (in hard tasks)

Human raters confirmed **higher clarity in size, position, color, and style.**

Task 2 – Exploration

RQ: How does BRICKIFY influence users' creative exploration when they start without a clear intent?

Task: create 3 storybook scenes about an owl's adventures.
Focus: creativity, consistency, and reuse of tokens. Metrics:

- Creativity Support Index
- Token usage




Task 2 - Results

Brickify shows strong support for creativity, effectively supports idea exploration, and is generally enjoyable to use.

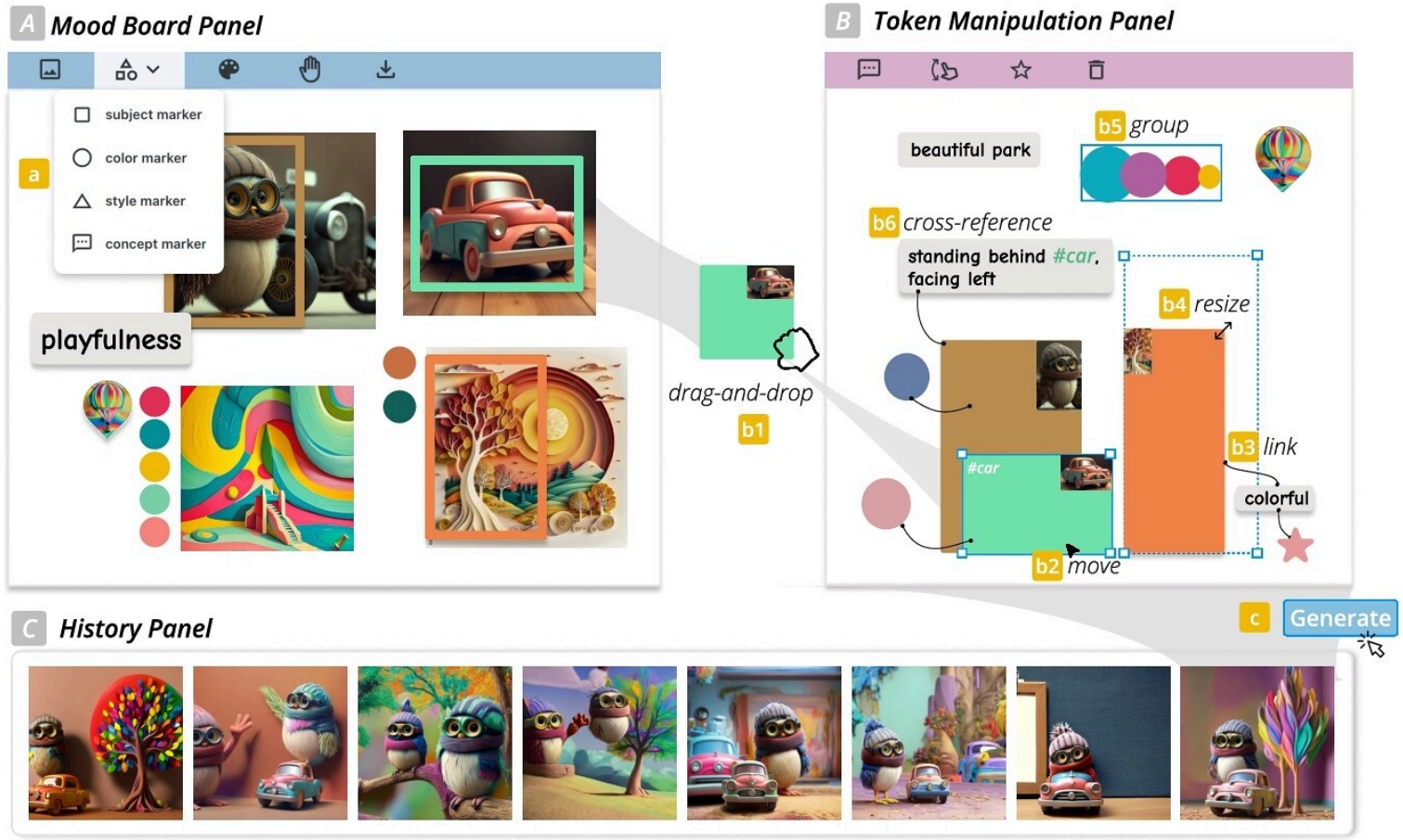
7. DISCUSSION

Limitations

- Visual lexicon extraction could be improved
- Inference and computation costs could hinder user experience
- Brickify might fail in describing unseen visuals beyond recombination
- Study results might not be generalizable for design novices

What about Multiple views and projection?

A *multiple view system* uses two or more distinct views to support the investigation of a single conceptual entity.



Thank You

Questions?
https://doi.org/10.1145/3706598.3714087