Investigating the Mismatch in Affective Evaluation and Design Processing Between Expert and Novice Product Designers on AI-generated Designs

Achmad Syarief[1], Yannes M Pasaribu[1], Yan Yan Sunarya [2], Muhammad Ridhwan[1] , Hanifah M Az-zahra [3]
[1] Research Group of Human and Industrial Product Design – FSRD ITB, 2 Research Group of Craft and Tradition – FSRD ITB, 3 Doctoral Student - FSRD ITB

Abstract

Generative AI enables rapid product design exploration, but visual appeal does not necessarily indicate technical feasibility. This mixed-methods study examined how 37 design students and 39 professional designers evaluated six AI-generated product designs. Quantitative analysis showed no significant overall differences between groups. However, qualitative analysis of 456 open-ended responses revealed different evaluation priorities. Students more frequently emphasized aesthetics, materials, and comfort, while professionals considered functionality, structural integrity, production feasibility, and market factors more often. These findings suggest that similar evaluation scores can conceal different perspectives, potentially shaped by educational and professional experiences. The study highlights the importance of integrating professional perspectives into design education through greater exposure to materials, construction, prototyping, and manufacturing practices.

Keyword: AI-generated design; design expertise; affective evaluation; analytical judgment; technical feasibility; product design.

Introduction

Generative AI enables rapid exploration of product forms, but visually appealing concepts may not always be functionally or technically feasible. Evaluating these designs requires consideration of aesthetics, usability, structure, and manufacturing possibilities. Differences in knowledge and experience between design students and professional designers may influence their evaluation priorities. This study investigates differences in affective responses, technical judgments, and evaluation focus between the two groups when assessing AI-generated product designs.

Method

This mixed-methods study involved 76 participants (37 design students and 39 professional designers) evaluating six AI-generated product designs: two chairs, two kettles, and two desk lamps. Participants completed five measures covering AI-generated impression, aesthetics, Good Design principles, adapted SAM responses, and perceived technical viability. They also provided spontaneous open-ended associations for each stimulus.

Quantitative data were analyzed using mixed ANOVA to examine group, stimulus, and interaction effects. A total of 456 open-ended responses were analyzed following Creswell's six-step qualitative approach to identify themes and differences in evaluation focus.



Fig 1. AI-generated product stimuli used in the study: chairs (STIM_01–02), kettles (STIM_03–04), and lamps (STIM_05–06)

	STIM_01 Chair	STIM_02 Chair	STIM_03 Kettle	STIM_04 Kettle	STIM_05 Lamp	STIM_06 Lamp
MIN – Minimal / simple	28	4	43	45	54	45
EST – Aesthetic / attractive	28	13	12	26	21	26
MOD – Modern / futuristic	12	17	18	21	12	18
REF – Design reference / icon	12	4	7	11	9	4
HGT – Warm / natural / traditional	46	0	0	33	0	0
DIN – Cold / stiff / monotonous	1	11	18	0	17	8
MAT – Material identification	30	4	22	33	1	8
ECO – Sustainable / eco-friendly	5	1	1	8	1	0
DEN – Product type & function	25	34	30	22	30	26
FUN – Functional / fit for purpose	9	18	25	14	30	9
NYM – Comfortable / ergonomic	32	64	1	1	3	0
KUG – Ease of use / ergonomics	5	8	11	8	12	21
PHM – Shape readability (easy to understand)	7	9	5	11	7	12
KKH – Sturdy / strong / durable	8	1	11	3	4	0
RKH – Compact / slim / not bulky	5	3	5	5	0	20
RMT – Complex / over-engineered	7	29	0	1	0	5
PRP – Proportion / balance / form harmony	7	4	4	3	4	9
GEN – Generic / conventional / already on the market	14	3	18	1	18	1
INO – Innovative / unique / experimental	4	9	1	4	0	13
PRD – Manufacturability & market readiness	5	4	4	3	7	3
HRG – Value / price / premium	3	5	3	3	1	1
KTX – Context & usage association	16	20	9	13	13	8
AI – AI trace / AI-like rendering	8	1	3	3	0	3

Fig 2. Qualitative Evaluation Patterns Across Six AI-Generated Design Stimuli (n = 76 per stimulus).

Results & Discussion

Quantitative analysis found no significant overall differences between students and professionals across five evaluation measures. However, scores varied significantly across stimuli, with a significant aesthetic difference for STIM_03, which professionals rated more positively than students.

The stimulus heatmap (Fig. 2) reveals contrasting perceptions: STIM_02 was associated with both comfort (64%) and complexity (29%), while STIM_06 raised concerns about usability (21%) and structural fragility (20%). These findings indicate that visually distinctive designs can evoke both positive impressions and practical concerns.The key finding emerged from the qualitative comparison (Fig. 3). Students more frequently emphasized materials (−7.4 pp), comfort (−4.8 pp), and aesthetics (−4.6 pp), whereas professionals more often mentioned design references (+4.4 pp), durability (+4.2 pp), functionality (+3.9 pp), and production/market considerations (+2.9 pp).

These patterns suggest different evaluative priorities despite similar overall scores. Students tended to describe perceptual qualities, while professionals more often connected visual features to practical design considerations. Such differences may reflect educational and professional experiences, although this relationship was not directly tested.

Conclusion

Design expertise may be reflected more clearly in what designers consider than in the scores they assign. Understanding design evaluation therefore requires examining qualitative judgments alongside numerical ratings.

These findings highlight the need for design education to incorporate professional perspectives, connecting aesthetic development with practical knowledge of materials, construction, prototyping, and manufacturing.

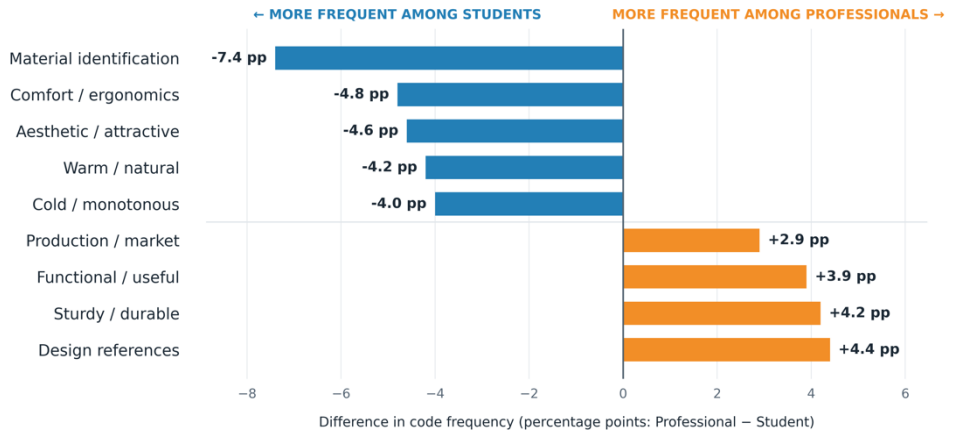


Fig 3. Differences of Focus Between Students and Professionals. Percentage-point differences in qualitative code frequencies.