

Visual Rightness of Text Overlay and Image Combinations in Social Media Communication

Stefania Farace (University of Bologna), Francisco Villarroel Ordenes (University of Bologna), Dennis Herhausen (Vrije Universiteit Amsterdam), Dhruv Grewal (Babson College), Ko de Ruyter (King's Business School)

Keywords: Salience, text overlays, dynamic images, visual rightness, social media engagement

A popular visual tactic to spark engagement that has emerged in social media communication is the use of text overlays (TOs) (Verlarde 2017), that is, words or phrases (beyond a brand trademark or logo) that are intentionally placed on a delimited area of the image included in a social media post. However, the fact that TOs can be integrated in several ways and in different images raises the question of when and how TOs increase consumer engagement with a social media composition.

We turn to visual rightness theory (Arnheim 1954; Locher 2003), which is based on the idea that the way visual elements are presented in a composition can feel intuitively right (Locher 2001) to understand the interplay between TOs and images in social media communication. We suggest that TOs and image must be combined carefully or else, they might interfere with each other and intuitively feel incorrect in terms of how the composition is being structurally executed (Affonso and Janiszewski 2022). We examine how three fundamental characteristics that are commonly used by practitioners to trigger salience, that is, the degree of dynamism, or implied motion in images, combined with the size and centrality of the TO, may affect consumer responses. Specifically, we argue that combining salient dynamic images with *large* TO positioned *centrally* has a negative effect on visual rightness and subsequently on consumer engagement.

A social media post where a large TO is combined with a highly dynamic image, could arguably feel intuitively incorrect decreasing perceptions of visual rightness for consumers, who have implicit preferences about how something should look to be visually pleasing. Specifically, the whole composition must be visually balanced in terms of salience that is conferred by both the dynamic image and the size of the TO to result in a “good” composition. However, if the TO becomes too large, such that the TO dominates the dynamic image, it can take away the positive feeling of visual rightness. We test our predictions in two online experiments.

METHOD

STUDY 1: LAB EXPERIMENTS ON TEXT OVERLAY EFFECTS

We randomly selected 20 different real tweets (10 static, 10 dynamic); none of these posts included TO, so we added them, manipulating their size and centrality, such that we generated 120 different social media posts. In both studies, we use a mixed-design approach, in which the image (dynamic vs. static) is a between-subjects factor, the TO size (small,

medium, large) or centrality (peripheral, somewhat central, central) are within-subjects factors.

TO Size. Using a sample of 145 Prolific participants ($M_{\text{age}} = 37.42$ years, 45.5% women) and panel regression, we find an inverted U-shape effect of TO size for dynamic images (main: $\beta = .291$, $p < .10$; squared: $\beta = -.188$, $p < .01$) but no significant effect for static images (main: $\beta = -.038$, $p = .83$; squared: $\beta = .015$, $p = .86$).

TO Centrality. Using a sample of 157 Prolific participants ($M_{\text{age}} = 39.44$ years, 49.7% women) and panel regression, we find that TO centrality has a negative effect for dynamic images ($\beta = -.127$, $p < .05$) but no significant effect for static images ($\beta = -.087$, $p = .22$).

STUDY 2: MEDIATING EFFECT OF VISUAL RIGHTNESS

In this study, we test our prediction that visual rightness drives the effect of image and TO salience. In each study, we use a three-factor (TO size: small, medium, large) (TO centrality: peripheral, somewhat central, central) within-subjects design that features three different product images (in total, nine different social media ads).

TO Size. Using multilevel analysis on a sample of 113 participants ($M_{\text{age}} = 37.09$ years, 46.7% women), we find that visual rightness mediates the inverted U-shaped effect of TO size on preference ($\beta = -.252$, $p < .001$).

Study 3b: TO Centrality. Using multilevel analysis on a sample of 116 participants ($M_{\text{age}} = 40.52$ years, 50% women), we again find that visual rightness mediates the negative effect of TO centrality on preference ($\beta = -.187$, $p < .001$).

GENERAL DISCUSSION

Across all studies, we consistently find that the interplay of desirable elements in multimodal social media communication can backfire. If a TO is too small or too big and placed at the center of a dynamic image, it mitigates the positive effects on engagement. Visual rightness explains these effects of TO size and TO centrality on consumer responses for dynamic images. Even if they tend to be processed quickly, social media multimodal compositions evidently can evoke consistent visual aesthetics evaluations.

SELECTED REFERENCES

- Affonso, Felipe M., and Chris Janiszewski (2023), "Marketing by Design: The Influence of Perceptual Structure on Brand Performance," *Journal of Marketing*, 87(53), 736–754.
- Arnheim, Rudolf (1954), *Art and visual perception: A psychology of the creative eye*. Univ of California Press.
- Cian, Luca, Aradhna Krishna, and Ryan S. Elder (2014), "This Logo Moves Me: Dynamic Imagery from Static Images," *Journal of Marketing Research*, 51 (2), 184–197.
- Hagtvedt, Henrik, Vanessa M. Patrick, and Reidar Hagtvedt (2008), "The Perception and Evaluation of Visual Art," *Empirical Studies of the Arts*, 26 (2), 197–218.
- Locher, Paul J. (2003), "An Empirical Investigation of the Visual Rightness Theory of Picture Perception," *Acta Psychologica*, 114 (2), 147–164.
- Pieters, Rik and Michel Wedel (2004), "Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects," *Journal of Marketing*, 68 (2), 36–50.