Participants’ Strategy in Crowd-Based Design Contests – A Prospect Theory Perspective

Completed Research Paper

Tat Koon Koh
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon Hong Kong
koh@ust.hk

Abstract

Using crowd-based contests to acquire creative designs is increasingly popular. In this study, I examine how client-provided information in design contests affect participants’ strategy and contest submissions. In these contests, clients often provide examples of designs that they like, thereby signaling their design preference. Using prospect theory, I relate participants’ contest payoffs and cost considerations to the extent that they deviate their submissions from client-provided examples. Results from a banner ad design contest show that participants (i) deviate their designs more from clients’ examples when the examples’ quality is lower, and when the examples are relatively similar, and (ii) submit designs that are more distinctive in contests when they deviate from clients’ examples. These results indicate that participants do not always seek to win contests by aligning their submissions to clients’ preference. Instead, participants’ strategy also depends on other benefits of contest participation and their preference to minimize participation costs.

Keywords: Online contests, wisdom of the crowd, crowdsourcing, banner ads, prospect theory, experiment
Introduction

For some time, firms have been seeking ideas and solutions from outside sources such as their key customers (Lilien et al., 2002; von Hippel, 2005). Recent developments in information technology and, particularly, crowdsourcing provide firms with another source of external inputs — anyone in the crowd. Crowdsourcing connects firms and solvers by facilitating the aggregation of solvers and the dissemination of firms’ problems (Jeppesen and Lakhani, 2010). Today, firms can use the crowd for a variety of activities, such as micro-tasks and software coding. The context of interest in this study is the crowd-based contests through which firms source for graphic designs such as banner ads. In these design contests, firms broadcast their design requirements to a large and diverse group of individuals, and allow anyone to submit their work. Firms then select and pay for design submissions that they want to acquire.

Design contests provide firms with various benefits. These contests allow firms to acquire ad designs within their budget. Firms can also receive a large number of design submissions from the crowd, thereby having more options to choose from. Furthermore, since firms only pay for those designs that they like, the mechanisms of design contests help firms reduce information asymmetry when sourcing for ad designs. Thus, by making it easier and more affordable to acquire ad designs, these contests are likely to contribute to the growth in online ad activities and spending. However, to maximize the benefits from using contests, firms need to know how they can influence contest participants to submit effective ad designs. To this end, research shows that designs that are more distinctive when compared to other designs in a set of alternatives tend to perform better in terms of ad recall, click-through rate, consumers’ attitude towards the ads, and consumers’ purchase intention (Brown 2002; Heiser et al., 2008; Li and Bukovac, 1999; Sundar and Kim, 2005). In contests for ad designs, within-contest design distinctiveness, or how different a submission is from all other submissions in a contest (Koh, 2013), is a key design outcome that firms should therefore consider.

In this study, I address two research questions that relate to within-contest design distinctiveness: how can firms shape participants’ strategy, and how does participants’ strategy affect distinctiveness of their design submissions in contests? Specifically, I model contest participants’ strategy using prospect theory (Kahneman and Tversky, 1979); this theory is appropriate for investigating decision-making under risk, which is the case in crowd-based contests given the uncertainties of contest outcomes for participants. Based on prospect theory, I consider how participants’ strategy is influenced by (i) benefits of winning and not winning contests, and (ii) loss aversion, where “losses loom larger than gains” for most individuals (Kahneman and Tversky, 1979: 279). By accounting for different contest outcomes and participants’ value function, this study challenges the assumption that participants’ strategy is mainly determined by their desire to win, and that a key objective is to satisfy firms’ taste (e.g., Terwiesch and Xu, 2008). This study also relates to research that examines impacts of design-level information (e.g., feedback for individual submissions during the contests) and participants’ behaviors (e.g., time of entry into the contests) in contests (Bockstedt et al., 2011; Wooten and Ulrich, 2011). At a broader level, this study further contributes to the application of prospect theory in decision making in IS research. Extant studies have used prospect theory to examine issues such as software project escalation (Keil et al., 2000) and technology usage (Brown et al., 2012; Venkatesh and Goyal, 2010). These studies focus on how the value function of losses is steeper than that of gains, and discuss the effects of loss aversion on individuals’ behavior. In this study, besides considering contest participants’ loss function, I also incorporate various types of potential payoffs to participants in their motivations and strategy.

Research Context

Crowd-Based Design Contests

Here is a typical process in a design contest for banner ads (and other creative designs such as corporate logos). Before launching a contest, the contest client decides on the contest duration and monetary award for winning designs. Contests have relatively short durations (e.g., between 3 to 14 days), and most design marketplaces provide recommendations for minimum awards. Next, the client provides a project brief that describes what he looks for in the banner design (Figure 1). He can state the design specifications and requirements, provide information about his business and target audience, and/or show some examples
of designs that he likes. Once the contest is launched, participants can submit their design entries anytime within the specified time frame. During the contest, the client can give feedback for some or all submitted designs at his discretion (e.g., Wooten and Ulrich, 2011). At the end of the contest, the client chooses designs that he wishes to acquire, and awards prizes to winning participants.

This figure shows a partial project brief in a design contest hosted on www.crowdspring.com. The client provided information of his design preference, and listed examples of banner design styles that he liked.

Figure 1. A Project Brief in a Design Contest

Participants’ Costs and Payoffs in Contests

Participants incur certain costs when they take part in design contests. These costs include their time and effort in coming up with design concepts and ideas, which involve the exploration and exploitation of solution space (e.g., March, 1991). In creative design tasks, exploration involves looking for major new design concepts that are applicable to the project objectives (e.g., Dorst and Cross, 2001; Dow et al., 2010). As participants explore the design solution space, they can discover unique concepts that few other participants may consider in design contests. In contrast, exploitation involves applying patches or refining existing design concepts to achieve slightly improved versions (Ball et al., 1994). Typically, exploitation is comparatively less costly than exploration in design projects; creating multiple designs based on a particular concept requires less resources and effort as compared to generating the same number of designs using alternative and distinctive concepts. It is crucial to relate contest participants’ strategy to their exploration and exploitation of design solution space. Insufficient search in the solution space could cause participants to become fixated on solutions too early, whereas excessive exploration may result in them having to spend more time and effort managing a diverse set of designs rather than improving on specific alternatives (Fricke, 1996).

Since participants cannot recoup their participation costs unless they win the contests, it is possible that nothing matters more to participants than having their submissions selected by clients so as to win the monetary awards (e.g., Sun et al., 2012). In addition, the prestige of winning and being highlighted as winners on contest platforms can improve participants’ status within the contest community, and provide valuable ego rewards to them (e.g., Fang and Neufeld, 2009; Sun et al., 2012). Thus, the desire to beat the competition for prizes and recognitions can be a key motivational force for participants, and extant research tends to focus on winning outcomes in contest participation (e.g., Bockstedet et al., 2011; Jeppesen and Lakhani, 2010; Yang et al., 2008). As and when winning matters, participants are likely to adopt a “client is king” attitude and strive to satisfy clients’ needs and tastes. They would pay attention to signals of clients’ preference to determine what constitutes good solutions (at least as perceived by the clients) (Terwiesch and Xu, 2008). An example of such signals is clients’ feedbacks to participants’ submissions in contests. These feedbacks help participants discover clients’ quality function (or taste),
and adjust their design strategy accordingly (Bockstedt et al., 2011; Wooten and Ulrich, 2011). By aligning their submissions to clients' preference, participants could increase their chances of winning the contests.

However, winning monetary prizes and prestige are not the sole benefits that participants can derive from design contests, and satisfying clients' preference may not always be their paramount goal. Because design contests offer participants practical learning opportunities through real-world projects, these contests also help participants gain design experience and sharpen their skills. Moreover, participants can use their submissions in the contests to enhance their design portfolios, and showcase these works to prospective clients in future projects, crowd-based ones or otherwise. I term these improvements to participants' human capital and profiles as the marketability impacts of contest participation. We see similar effects in other IS contexts such as open source software (OSS) projects. For example, although programmers may not receive any direct payoff for their efforts in OSS projects, they can improve their programming skills and signal to potential employers their talent through the OSS code that they contribute (Fang and Neufeld, 2009; Lakhani and Wolf, 2005; Roberts et al., 2006).

Theory and Hypotheses Development

Using prospect theory (Kahneman and Tversky, 1979), I integrate different participants' payoffs and their cost considerations to examine participants' strategy. Prospect theory has been applied to study individuals' decision-making in situations with probabilistic outcomes, similar to those that participants face in contests (where winning is uncertain). Specifically, in this case, contest participants' strategy depends on the different possible outcomes and corresponding payoffs. If their submissions were selected by contest clients, participants would receive contest prizes and winning prestige, as well as enhance their marketability. Alternatively, even if their submissions were not selected, participants could still improve their marketability by having participated in the contests. Based on prospect theory, I incorporate these two scenarios into the value of participants' contest participation prospects, \( V(\cdot) \):

\[
\begin{align*}
\text{(1)} & \quad V(x, p; y, 1-p) = v(y) + \pi(p) * [v(x) - v(y)] \\
\text{(2)} & \quad V(x, p; y, 1-p) = (1 - \pi(p)) * v(y) + \pi(p) * v(x)
\end{align*}
\]

In Equation 1, \( v() \) denotes the subjective value of an outcome, \( x \) is the outcome of winning the contest, \( y \) is the outcome of not winning the contest, \( p \) is the contest winning probability, and \( \pi() \) is a decision weight that indicates the importance of winning in the prospects (where \( 0 < \pi(p) < 1 \)) (Kahneman and Tversky, 1979). Due to the additional benefit of receiving prizes and prestige when participants win the contests, I assume \( v(x) > v(y) \). In the contest participation prospects, \( v(y) \) is the riskless portion, while \( v(x) - v(y) \) represents the risky portion. Equation 1 can also be rearranged into:

\[
\begin{align*}
\text{(2)} & \quad V(x, p; y, 1-p) = (1 - \pi(p)) * v(y) + \pi(p) * v(x)
\end{align*}
\]

From Equations 1 and 2, we see that participants should favor situations and strategies that increase \( p \), \( v(x) \), and/or \( v(y) \), ceteris paribus, so as to improve the value of their contest participation prospects. For example, participants would prefer projects with fewer competitors (higher \( p \)) and higher monetary award (higher \( v(x) \)) (Yang et al., 2008). In addition, because participants have limited resources and could not recover incurred costs unless they win the contests, they tend to be loss averse (Kahneman and Tversky, 1979). An implication of loss aversion is that participants would minimize their contest participation costs whenever possible. As such, when they are searching for design solutions, participants are likely to exploit the solution space whenever possible in contests, and explore only when necessary.

Given the participants' prospects and their loss aversion, I next examine how client-provided information affect participants' strategy and contest outcomes. The strategy of interest is design deviation, or the extent to which participants deviate their submissions from client-provided design examples in a contest. In design projects, clients often show examples of designs that they like (see Figure 1), and these examples could serve as a starting point for participants in their search for design solutions. Design examples provide inspirations for creative works, and help participants assess the originality of their ideas and identify flaws or limitations to avoid (Herring et al., 2009). Participants can also learn how others have approached a design problem by referring to existing designs. Moreover, client-provided examples give indications of clients' preferences, which could affect the selection of winning submissions in contests (Terwiesch and Xu, 2008).
In the next section, I relate contest participants’ payoffs and cost considerations to their exploration/exploitation of solution space to develop hypotheses concerning (i) the extent to which participants deviate from clients’ examples in contests, and (ii) how design deviation affects distinctiveness of participants’ submissions in the contests. Table 1 lists the main constructs in this study, and Figure 2 shows the research model.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Deviation</td>
<td>The extent to which a design differs from client-provided design examples in a contest.</td>
</tr>
<tr>
<td>Quality of Client-Provided Design Examples</td>
<td>The trait of clients’ examples in terms of their attractiveness and appropriateness for the design project.</td>
</tr>
<tr>
<td>Design Variability of Client-Provided Design Examples</td>
<td>The dissimilarity among clients’ examples in terms of their design concepts.</td>
</tr>
<tr>
<td>Within-Contest Design Distinctiveness</td>
<td>The extent to which a design differs from all other designs in a contest.</td>
</tr>
<tr>
<td>Design Divergence</td>
<td>The extent to which a design differs from other designs submitted by the same participant in the contest.</td>
</tr>
</tbody>
</table>

Table 1. Construct Definition

**Figure 2. Research Model**

**Impact of Quality of Client-Provided Design Examples on Design Deviation**

As clients like to receive high quality submissions that are attractive and appropriate for their needs in design projects, they usually provide design examples that they think reflect such characteristics. Although clients would not intentionally show low quality examples, the examples that they choose could be perceived as such by some contest participants. This is because the evaluation of design concepts, problems, and solutions by designers and non-designers could and often differ. Due to their design training and experience, some participants develop certain guiding principles or problem paradigms that affect how they approach design problems and develop solutions (Darke, 1979; Lloyd and Scott, 1994). Thus, clients who lack such background may not evaluate design concepts the way participants do.

If we assume participants’ key objective is to win contests, we would expect design deviation to be independent from the quality of client-provided design examples. As these examples reflect clients’ preference, participants could better satisfy clients and increase their winning chances by using concepts from these examples in their design submissions, regardless of the examples’ quality. On the contrary,
based on the prospect theory, the quality of clients’ examples should affect design deviation. When clients use high quality examples that are attractive and appropriate for the project, participants are likely to explore less of the solution space since they are generally loss averse and prefer to minimize their costs. Instead, they could build upon and use design concepts from these examples, causing their submissions to deviate from clients’ examples to a lesser extent.

On the other hand, when clients’ examples of low quality, participants are less likely to use concepts from these examples. This is because borrowing from inferior examples may not enhance the participants’ experience or skills, and could also hurt the quality of their portfolio (i.e., lower \(v(x)\) and \(v(y)\)). Moreover, by submitting unattractive designs, participants may also lower their chances of winning (lower \(p\)) even if these submissions are consistent with the stated preferences of clients. Hence, when clients provide low quality design examples, participants are likely to explore more and search other design concepts for ideas and inspirations. Consequently, their submissions should deviate more from the clients’ examples. Consistent with prospect theory, I thus expect participants to deviate less (more) from clients’ examples when the examples are of high (low) quality.

**H1: Quality of client-provided design examples is negatively related to design deviation.**

**Impact of Design Variability of Client-Provided Design Examples on Design Deviation**

Another potential factor of design deviation is the design variability of clients’ examples, or the degree to which the examples differ in design concepts such as layout and content. In design projects, clients may provide multiple design examples that they like. These examples could be relatively similar (low design variability) in some cases, and highly varied (high design variability) in other cases. The expected impact of design variability of clients’ examples on design deviation depends on how similarity or dissimilarity among the examples affect (i) the perceived specificity of clients’ design preference, and (ii) the perceived design benchmarks that participants aim to outperform.

When design variability of clients’ examples is low in contests, clients’ design preference could be perceived to be highly specific. To increase their winning chances (\(p\)) and reduce their participation costs, participants may restrict their exploration of the solution space and concentrate on design concepts in clients’ examples. Thus, when the emphasis is on winning the contests, design deviation should be positively related to design variability of clients’ examples, where participants would deviate less (more) from clients’ examples when the examples are relatively similar (different).

However, given the creative nature of design tasks, satisfying clients’ design preferences may not be the sole objective for participants. Many designers (such as those who take part in design contests) consider themselves as creative artists, and often set extremely high design standards and goals for themselves (Cross, 2003; Lawson, 1994). They also have personal design preferences and principles that guide their creative work (Darke, 1979; Lloyd and Scott, 1994). Many of them aim to develop creative solutions that are new to the market, and take pride in showcasing such novel works in their portfolios. Thus, while winning contests is important to participants, they also strive to come up with distinctive work as part of the process to enhance their marketability. We observe similar intrinsic motivations among participants in OSS project, where the high sense of personal creativity in OSS project participation is a key determinant of participants’ project involvements (Lakhani and Wolf, 2005).

From this perspective, a set of highly similar clients’ examples could then serve as a creative benchmark that participants want to distinguish their work from. When clients’ preference is perceived to be very specific, participants could anticipate competitors to actively incorporate design concepts from the clients’ examples. In order for their submissions to stand out amongst the competition, participants might be willing incur greater participation costs and explore the design solution space more widely for concepts that differ from those in the examples. On the one hand, this strategy may lead to designs that do not fully fit clients’ taste (lower \(p\)). On the other hand, by deviating from clients’ examples, participants would not be competing with a potentially large set of similar designs from their competitors (higher \(p\)). Although the overall effect on winning chances is ambiguous when participants deviate from clients’ examples in this case, doing so can help them create potentially distinctive designs and improve their marketability (higher \(v(x)\) and \(v(y)\)). Thus, based on prospect theory, it is possible that participants deviate more from clients’ examples when the design variability of the examples is relatively low:
**H2**: Design variability of client-provided design examples is negatively related to design deviation.

I also expect a stronger tendency for participants to deviate from similar clients’ examples as the number of examples increases. What clients like is more salient when many relatively similar examples are shown. In this case, participants might expect greater adherence to clients’ examples by competitors, causing them to deviate more from the examples than had clients provided fewer similar examples. Thus, I hypothesize the number of clients’ examples to moderate the relationship between design variability of the examples and design deviation:

**H3**: Design variability of client-provided design examples has a stronger (weaker) negative impact on design deviation when the number of examples is higher (lower).

Table 2 summarizes the expected relationships between attributes of clients’ examples and design deviation under different perspectives.

<table>
<thead>
<tr>
<th></th>
<th>“Client is King” Perspective</th>
<th>Prospect Theory Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contest participants’ goals</td>
<td>Participants’ primary objective is to win the contests for prizes and prestige. Satisfying clients’ design preference can improve participants’ winning chances.</td>
<td>Apart from winning prizes and prestige, participants can gain design experience and skills, and improve their portfolio by taking part in design contests.</td>
</tr>
<tr>
<td>Expected relationship between quality of client-provided design examples and design deviation (H1)</td>
<td>Design deviation is not related to quality of client-provided design examples: Since these examples are signals of clients’ preference, the optimal strategy for participants is to adhere to the examples – regardless of their quality – so as to increase their winning chances.</td>
<td>Design deviation is negatively related to quality of client-provided design examples: Participants would deviate more from low quality clients’ examples because borrowing design concepts from such example could lead to inferior submissions, which may hurt their winning chances and marketability.</td>
</tr>
<tr>
<td>Expected relationship between design variability of client-provided design examples and design deviation (H2, H3)</td>
<td>Design deviation is positively related to design variability of client-provided design examples: Lower design variability of the examples provides stronger signals of clients’ design preference, which lead to less design deviation.</td>
<td>Design deviation is negatively related to design variability of client-provided design examples: Lower design variability of the examples provides greater incentives for participants to deviate and create novel designs so as to stand out from the competition.</td>
</tr>
</tbody>
</table>

**Table 2. Alternative Perspectives of Participants’ Strategy**

**Impact of Design Deviation on Within-Contest Design Distinctiveness**

Existing studies indicate that, among alternative ad designs, those that are more distinctive perform better in advertising campaigns (e.g., Brown 2002; Heiser et al., 2008; Li and Bukovac, 1999; Sundar and Kim, 2005). These findings suggest that submissions with higher within-contest design distinctiveness are likely to be more effective. In this regard, how is within-contest design distinctiveness affected by the extent to which participants deviate from clients’ examples? I posit that contest participants are more likely to discover design concepts that relatively few competitors would consider when they deviate from clients’ examples and explore a wider area in the solution space. Using these alternative concepts should result in more unique submissions when compared to other submissions in the respective contests. Thus,
I hypothesize that participants’ submissions would be more distinctive within contests when participants deviate from clients’ examples.

**H4: Design deviation is positively related to within-contest design distinctiveness.**

### Method

#### Overview

I conducted a design contest where participants were asked to design banner ads to promote an online wedding photography directory. I recruited participants from various online communities for graphic designers, and invited all individuals, regardless of their design experience, to participate in this study. To make the experiment realistic to design contests, I did not compensate participants for participating in this study. Instead, participants who submitted the top three designs would each receive between US$250 and US$600. These amounts were consistent with the rates on various design contest platforms at the time of the study. Individuals who wished to participate in this study could pre-register by providing their names and email addresses.

I emailed the registrants once the design contest was launched. Registrants first answered a pre-contest survey, where I asked them design- and contest-related questions. After they completed the survey, they received login passwords for the design contest platform that I developed. Once they logged into the platform, they viewed the project brief that described the wedding photography directory and its target audience. They were asked to submit ad designs that are attractive, and would achieve high ad recognition performance and click-through rate. In the project brief, they also saw some examples of online banners that were selected prior to the experiment, which I will discuss in the next section.

I provided participants a logo for the online directory, and ten photos that they could include in their ad designs. These photos showed different wedding-related images such as a bride and/or groom (in various poses and different settings), wedding bouquet, and wedding gown. Due to legal and copyright concerns, participants must only use the photos that I provided in their designs. However, participants were free to create and use ad copy, such as tagline and phrases, in their submissions. I specified that the ad dimensions must be 300 (width) x 250 (height) pixels, and be less than 50kb in file size.

To minimize the impact of competition on designers' efforts during the contest (e.g., Boudreau et al., 2011), I employed a blinded contest structure where participants could not observe other participants’ submissions during the contest. This structure prevented participants from strategizing their designs based on their observations of the competition. I also did not indicate the number of participants who were taking part in this study, or the number of designs that had been submitted.

Participants had ten days to submit their designs through the platform, starting from the day they first logged on to the website. Participants could access the project brief, logo, and photos any time during the duration of the contest. They could also withdraw their design submissions any time before the contest ended. After the contest ended, I invited the participants to complete a post-contest survey with questions about the experimental design task.

### Experimental Design

The stimuli in this experiment were design examples that I showed to participants in the project brief. I planned to use participants’ evaluations of the examples to measure the examples’ quality and variability. To achieve different levels of examples’ design variability across participants, I used different categories of design examples, where examples within each category were relatively similar (low design variability), and examples across different categories were relatively dissimilar (high design variability). The four categories that I used in the experiment were (i) ads with collages, (ii) ads with wedding bouquet as focal point, (iii) ads with greenery background, and (iv) ads with top-and-bottom frames. Each category consisted of six banner ads promoting wedding photography services that a research assistant (blinded to the study) found on the Internet prior to the experiment.
Sample

I used an incomplete 4 x 2 experimental design, where I crossed number of examples (4 levels) with design variability of examples (low vs. high). Two conditions do not exist (high variability with zero or one example) so the design is an incomplete factorial design with six conditions. I included conditions with 0- and 1-example to examine if showing no examples or a single example would outperform or underperform showing multiple examples in terms of achieving distinctive designs. This additional analysis would give more insights into the relationships between clients’ examples and within-contest design distinctiveness.

Using a multi-step process, the system randomly assigned (i) participants to conditions, and (ii) stimuli (design examples) to participants when they logged on to the contest platform for the first time after completing the pre-contest survey. The system began by randomly choosing the number of design examples to assign (0, 1, 2, or 4 examples). If the participants were assigned to see one example, the system randomly selected an example from the pool of 24 stimulus ads. If they were assigned to see two or four examples, the system first randomly selected design variability of the examples (low or high). When the assigned variability was low, the system randomly selected an example category, and then randomly chose ads from that category. Alternatively, when the assigned variability was high, the system randomly chose ads from different example categories, but at most one ad from each category.

176 individuals completed the pre-contest survey, and 105 (59.7%) of them submitted at least one design during the contest. 385 submissions were received at the end of the contest, but 45 submissions were not usable for this study: 18 designs included photos that I did not provide and/or URL of other websites instead of the online directory that they were supposed to design the ads for, while the dimensions of 27 others were not within the specified width and/or height and could not be resized without removing key elements in the ads. Consequently, the sample consisted of 340 submissions from 99 participants.

Measures

**Level 1 (Submission-Level) Measures**

*Within-Contest Design Distinctiveness.* Submissions that are, on average, more different from the others in a contest are considered to be relatively distinctive (Koh, 2013). Therefore, I measured a submission’s distinctiveness by comparing and averaging its differences from all other submissions in the contest. In this experiment, calculating within-contest design distinctiveness requires the comparison of 57,630 submission pairs. I employed a 3-step approach to obtain these pairwise distances and assess the distinctiveness of individual submissions (Koh, 2013).

In Step 1, I *codified* each submission’s design attributes in terms of the color scheme, photos and logo, and text. These attributes are vital design elements in advertisements (Gorn et al. 1997; Hanssens and Weitz, 1980; Pieters and Wedel, 2004, 2012). I used a programming script to quantify the color scheme (in terms of the distribution of RGB decimal values) in each submission. Using an online interface, two coders measured the sizes occupied by specific photos, logo, and text in each submission. Based on the coding, I also obtained the number of photos that were used in each submission.

In Step 2, I *compared* submission pairs over two sub-stages to obtain the pairwise distances for all the submissions. In the first sub-stage, I calculated the differences in various design attributes (color schemes, specific photos used, number of photos, size of photos, size of logo, and size of text-area) for all submission pairs using the obtained values in Step 1. In the second sub-stage, I used a sub-sample of

---

1 Because examples in the high design variability conditions have to be relatively different from each other, the number of alternative categories limited the maximum number of examples I could show to participants. In addition, I did not include conditions for 3-high variability and 3-low variability examples. I only needed the extreme values (in this case, 2 and 4 examples, and high and low variability) to test linear interactions between quantity and design variability of examples (McClelland, 1997).

2 RGB triple represents the amount of red (R), green (G), and blue (B) that a pixel has. The values for each RGB component range from 0 to 255. For example, black has rgb(0,0,0) whereas white has rgb(255,255,255).
submissions to estimate how differences in the respective design attributes affect pairwise distances between submissions. To do this, I randomly selected 74 submissions (2,701 submission pairs) from the sample, and recruited raters to evaluate the similarity/dissimilarity of these submissions using the Spatial Arrangement Method (SpAM) (Goldstone, 1994; Hout et al., 2013). SpAM is a fast and efficient way to collect similarity/dissimilarity data when a large number of stimuli are involved, and the results using SpAM are comparable to those using traditional pairwise comparisons (see Hout et al., 2013). Using an online interface, five raters arranged sets of six randomly chosen submissions from the sub-sample; raters were instructed to arrange the submissions such that similar (dissimilar) ones should be placed closer to (further from) one another. Based on the submissions’ coordinates in the online interface for each set, I calculated the Euclidean distances for the various submission pairs, where longer distances indicate greater perceived dissimilarity between submissions. Each submission pair in the sub-sample was evaluated by at least three raters, and I averaged the pair’s Euclidean distances across all raters.

I then regressed the averaged Euclidean distances on the attribute differences that were calculated during the first sub-stage in this step; the estimated coefficients in this regression represents impacts of respective attribute differences on the perceived difference between two submissions. Differences in all attributes, except text-area sizes, are significantly and positively related to pairwise distances ($p < .001$ for differences in color schemes, specific photos used, number of photos, size of photos, and size of logo; $p > .10$ for difference in size of text-area).

Using the regression coefficients and attribute differences, I calculated the pairwise distances for all 57,630 submission pairs in the sample. To validate the robustness of the regression-based pairwise distance estimates, I randomly selected 10 pair of submissions within the 95% confidence interval at (i) every 10th percentile (i.e., 10th, 20th, ..., 90th) in terms of the computed pairwise distance, and (ii) the 1st and 99th percentile. I also selected five pairs of submissions among those with the highest computed distances, and five pairs among those with the lowest computed distances. I asked users on Amazon Mechanical Turk (AMT) to rate the similarity of these 120 submission pairs on a 7-point scale. The dissimilarity scores were obtained by averaging the reverse-coded ratings. I found a high correlation between the AMT dissimilarity scores and the regression-based pairwise distances ($r = .82$, $p < .001$), which validates the estimates.

Finally, in Step 3, I calculated within-contest design distinctiveness of each submission by averaging its regression-based pairwise distances with all other submissions in the sample.

**Design Deviation (Submission-Level).** I recruited raters on AMT to compare every example assigned to individual participants with each of the respective participants’ submissions. Raters were randomly assigned pairs of examples and submissions, and they evaluated the similarity of each pair on a 7-point scale. For each example-submission pair, I reverse-coded and averaged all its ratings to obtain its dissimilarity score. From here, I averaged all the dissimilarity scores of a particular submission to determine its design deviation score.\(^3\)

---

\(^3\) Apart from average dissimilarity scores, I considered alternative measures for design deviation, such as minimum or maximum dissimilarity scores of each submission. A reason that favors the use of average scores is that it accounts for all the assigned examples that participants saw. In contrast, minimum/maximum scores implicitly focus on difference between a submission and a specific example, and ignore how that submission differs from other examples assigned. Therefore, average scores provide richer information than minimum/maximum scores. However, average dissimilarity scores could suffer from certain statistical confound, especially if participants systematically chose one assigned example to base each of their submissions on. For example, holding the number of assigned examples constant, the mean of average dissimilarity scores could be systematically lower in a low design variability condition than in a high design variability condition, since the examples are relatively similar to each other in the former condition. Likewise, between the two high design variability conditions, the mean of average dissimilarity scores could be systematically higher in the 4-example condition than that in the 2-example condition, as the submission is different from more examples in the 4-examples condition. To check whether it is appropriate to use average dissimilarity scores as a measure for design deviation, I examined the average dissimilarity scores across experimental conditions. I found no significant differences in the mean of average dissimilarity scores between (i) 2-Example/Low Design Variability and 2-
**Design Divergence.** Participants could submit multiple designs in the experiment. Design divergence is the extent to which a design submission differs from the others by the same participant (Dow et al., 2010). This measure thus requires the pairwise distances between submissions by the same participants. Since I estimated the distances of all submission pairs when computing within-contest design distinctiveness (see second step in the 3-step approach), I used these estimates to calculate every submission’s design divergence by averaging its pairwise distances with other submissions from the same participant.

**Level 2 (Participant-Level) Measures**

*Design Deviation (Participant-Level).* I averaged the design deviation of all submissions by the respective participants to compute the person-level design deviation (e.g., Raudenbush and Bryk, 2002). This variable captures between-participant effects of deviating designs from assigned examples.

*Number of Examples.* I recorded the number of examples (0, 1, 2, or 4) that were assigned to the participants during the experiment.

*Quality of Examples.* In the post-contest survey, I asked participants who were assigned at least one example to rate the quality of the example(s). The participants evaluated the (i) appropriateness and (ii) attractiveness of each assigned example for the design project on a 7-point scale (Cronbach’s alpha = .89). I averaged all participants’ ratings for each example to improve the accuracy of the quality measure. To determine the examples’ quality for individual participants, I averaged the quality ratings of all examples that they were assigned to in the experiment.

*Design Variability of Examples.* Participants who saw multiple examples rated the similarity of those examples in the post-contest survey. They evaluated each pair of examples in terms of their (i) overall similarity, (ii) layout, and (iii) images on a 7-point scale (extremely dissimilar to extremely similar) (Cronbach’s alpha = .83). I reverse-coded their ratings to compute the dissimilarity between each pair of examples. I then averaged the dissimilarity ratings across participants for every pair of examples. To measure design variability of examples for individual participants, I averaged the dissimilarity ratings for relevant pairs of assigned examples.

*Perceived Marketability Impacts.* The importance of winning contests (or the decision weight, \(\pi()\), in Equations 1 and 2) could vary across participants and shape individual participants’ strategy in contests. To account for participant heterogeneity, I controlled for the perceived effects of contests participation on participants’ marketability. I asked participants the extent to which participating in design contests impact their (i) design experience, (ii) design skills, and (iii) design portfolio (“extremely negative impact… extremely positive impact”; Cronbach alpha = .94). Compared to participants who felt taking part in contests could contribute to their marketability, those who felt contest participation had negative marketability impacts were likely to focus more on their winning chances in their design strategy.

*Domain Experience.* As participants were recruited from the field, it is necessary to account for their experience in different categories of graphic design projects (e.g., Boons et al., 2013; Jeppesen and Lakhani, 2010). In the pre-contest survey, participants reported the number of projects in various categories in which they participated in the past two years. I calculated the ratio of wedding-related projects to all other types of projects to control for participants’ experience in the particular project domain in the experiment.

**Results and Analyses**

In the main analyses, I used observations in the multiple-example conditions since one of the research questions is how participants’ strategy is affected by design variability of examples (which does not exist in 0- and 1-example conditions). However, I conducted additional analyses involving all experimental conditions to examine the relationships between the client-provided design examples and within-contest Example/High Design Variability, (ii) 4-Example/Low Design Variability and 4-Example/High Design Variability, and (iii) 2-Example/High Design Variability and 4-Example/High Design Variability (\(p > .10\) in all cases for the null hypothesis that the mean scores between the two conditions are equal). These results provide support for using average dissimilarity scores as a measure of design deviation.
design distinctiveness. Given the multi-level structure of the data, I used Hierarchical Linear Modeling in the analyses. To facilitate the reporting of the estimates, I scaled the within-contest design distinctiveness variable by multiplying the score by a constant (100). Table 3 shows the descriptive statistics and correlation matrix.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Distinctiveness</td>
<td>1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Marketability Impacts</td>
<td>2</td>
<td>0.12</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain Experience</td>
<td>2</td>
<td>-0.02</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Divergence</td>
<td>1</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Deviation (Submission-Level)</td>
<td>1</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.17</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Deviation (Participant-Level)</td>
<td>2</td>
<td>0.19</td>
<td>0.21</td>
<td>-0.22</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Examples</td>
<td>2</td>
<td>0.06</td>
<td>-0.09</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Examples</td>
<td>2</td>
<td>-0.02</td>
<td>0.37</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.15</td>
<td>-0.30</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Design Variability of Examples</td>
<td>2</td>
<td>-0.01</td>
<td>0.34</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.09</td>
<td>0.11</td>
<td>-0.10</td>
<td>1.00</td>
</tr>
</tbody>
</table>

| Mean   | -0.79 | 0.00   | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    |

| Std. Dev. | 8.60  | 1.13   | 0.15    | 0.11    | 0.62    | 0.50    | 0.99    | 0.42    | 1.14    |

I scaled the original score of item 1 by a constant (100), and mean-centered items 2-9.

Table 3: Descriptive Statistics and Correlation Matrix

Table 4 shows the results. In Model 1, the baseline model, I included participants' perceptions of marketability impacts and their domain experience. Next, in Model 2, I included the quantity, quality, and design variability of examples. These variables explained 22.2% of the variation in design deviation. Design deviation was negatively affected by examples' quality ($\beta = -0.47$, $p < .01$) and design variability ($\beta = -0.13$, $p < .05$), supporting H1 and H2. In Model 3, I added the interaction between the number and design variability of examples. This interaction is not significant ($\beta = 0.02$, $p > .10$), and H3 is not supported.

<table>
<thead>
<tr>
<th>DV: Design Deviation $^\wedge$</th>
<th>Model 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.43 $^\wedge$ (0.07)</td>
<td>5.43 $^\wedge$ (0.07)</td>
<td>5.43 $^\wedge$ (0.07)</td>
</tr>
<tr>
<td>Perceived Marketability Impacts</td>
<td>0.06 (0.06)</td>
<td>0.18 $^\wedge$ (0.07)</td>
<td>0.17 $^\wedge$ (0.07)</td>
</tr>
<tr>
<td>Domain Experience</td>
<td>-0.67 (0.44)</td>
<td>-0.74 $^\wedge$ (0.42)</td>
<td>-0.72 $^\wedge$ (0.42)</td>
</tr>
<tr>
<td>Number of Examples</td>
<td>0.03 (0.07)</td>
<td>0.03 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Quality of Examples</td>
<td>-0.47 $^\wedge$ (0.18)</td>
<td>-0.45 (0.19)</td>
<td></td>
</tr>
<tr>
<td>Design Variability of Examples</td>
<td>-0.13 $^\wedge$ (0.06)</td>
<td>-0.13 $^\wedge$ (0.06)</td>
<td></td>
</tr>
<tr>
<td>Number X Design Variability of Examples</td>
<td>0.02 (0.07)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>var( Constant )</td>
<td>0.09 (0.05)</td>
<td>0.07 (0.04)</td>
<td>0.07 (0.04)</td>
</tr>
<tr>
<td>var( Residual )</td>
<td>0.53 (0.06)</td>
<td>0.51 (0.06)</td>
<td>0.51 (0.06)</td>
</tr>
<tr>
<td>Wald statistics</td>
<td>3.23</td>
<td>14.50 $^\wedge$</td>
<td>14.62 $^\wedge$</td>
</tr>
</tbody>
</table>

$^\wedge$ I used the original (non-mean-centered) values of design deviation in the analyses here. Standard errors in parentheses. $^p < .10$  $^*p < .05$  $^{**}p < .01$

Table 4. Impacts of Design Examples on Design Deviation
Next, I examined how the degree to which participants’ submissions differed from assigned examples affected within-contest design distinctiveness. Table 5 shows the results. Model 1 is the baseline model that includes perceived marketability impacts, participants’ domain experience, and design divergence of each design. In Model 2, I added design deviation at the submission- and participant-levels. The change in the variance component of the intercept between the models indicates that design deviation explained 12.2% of the variation in within-contest design distinctiveness. Design distinctiveness was positively affected by design deviation at the participant-level ($\beta = 2.82, p < .05$), supporting H4.

<table>
<thead>
<tr>
<th>DV: Within-Contest Design Distinctiveness</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.00 (0.78)</td>
<td>-0.84 (0.76)</td>
</tr>
<tr>
<td>Perceived Marketability Impacts</td>
<td>1.15* (0.68)</td>
<td>0.96 (0.66)</td>
</tr>
<tr>
<td>Domain Experience</td>
<td>-0.01 (4.95)</td>
<td>1.87 (4.87)</td>
</tr>
<tr>
<td>Design Divergence</td>
<td>27.65*** (8.63)</td>
<td>27.79*** (8.71)</td>
</tr>
<tr>
<td>Design Deviation (Submission-Level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Deviation (Participant-Level)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Impacts of Design Deviation on Design Distinctiveness

Additional Analyses

Instead of providing multiple design examples, could showing no or one example improve/hurt design distinctiveness in the contests? To address this question, I conducted two analyses to check if within-contest design distinctiveness is significantly different across the experimental conditions.

First, I regressed within-contest design distinctiveness on five condition dummies (0-Example, 2-Example/Low Design Variability, 2-Example/High Design Variability, 4-Example/Low Design Variability, and 4-Example/High Design Variability), with the 1-Example condition as the reference group. The overall F-test is not significant, and the model explains a very small proportion of the variance in within-contest design distinctiveness ($F = 1.93, p = .089, R^2 = .029$). Hence there is no evidence that within-contest design distinctiveness differ across the six experimental conditions.

Second, I re-categorized the four groups with multiple examples into two conditions based on design variability (low or high). I then regressed within-contest design distinctiveness on three condition dummies (0-Example, Low Design Variability, and High Design Variability), again using the 1-Example condition as the reference category. The overall F-test is not significant ($F = 1.99, p > .10, R^2 = .016$), indicating there were no systematic differences in within-contest design distinctiveness across the conditions.

Results from the main and additional analyses thus suggest that within-contest design distinctiveness is not affected simply by whether clients provide examples or how varied the examples are. Rather, within-contest design distinctiveness depends on the extent to which participants deviate from clients’ examples, which in turn is affected by factors such as quality and design variability of the examples.
Discussion

Using prospect theory (Kahneman and Tversky, 1979), I related contest participants’ strategy to the probability of contests winning ($p$), the subjective value of the winning outcome ($v(x)$), and the subjective value of the non-winning outcome ($v(y)$). I assume a positive $v(y)$ by accounting for improvements to participants’ marketability by taking part in contests. This riskless portion of contest participation prospects has not been sufficiently emphasized in existing research on crowd-based contests. In contrast, a common assumption is that winning contests is a key, if not the only, goal for participants. According to this “client is king” perspective, $v(y)$ is null, and participants’ strategy depends mainly on $p$ and $v(x)$ (e.g., Boudreau et al., 2011; Yang et al., 2008). However, the results in this study indicate that these two determinants, though important, may not fully explain participants’ strategy in crowd-based contest.

The perspective based on prospect theory that I employed in this study provides insights into why participants’ behaviors could appear counter-intuitive in contests. For example, they adopt strategy that may lower their winning chances, such as not incorporating design concepts from certain types of clients’ examples (H1), or deviating from the examples when they are highly similar (H2). These strategies could arise because participants can also improve their prospects by enhancing their marketability for future projects, even when they do not win the contests. Yet, at least in the case of ad design contests, such prospects-maximizing behaviors by participants are not necessarily bad for contest clients as they might receive more distinctive designs (H4), which could perform better in ad campaigns (see Brown 2002; Heiser et al., 2008; Li and Bukovac, 1999; Sundar and Kim, 2005).

As this study demonstrates, prospect theory provides a framework for us to examine the effects of contest mechanisms on participants’ strategy and creative processes. By looking at how a contest feature affects $p$, $v(x)$, and $v(y)$, we can make theory-driven predictions about the feature’s impact on participants’ behaviors. I believe this framework is particularly useful for those features that affect multiple aspects of contest participation prospects. For instance, to understand how the quality of competition could affect participants’ decisions to participate in a contest, we can look at its effects on $p$, $v(x)$, and $v(y)$. On the one hand, high quality competition is likely to decrease $p$. On the other hand, such competition could stimulate participants to create better work and improve their skills and portfolios, thereby increasing $v(x)$ and $v(y)$. Hence, the presence of high quality competition may not always drive participants away from a contest. As this abstract example shows, applying prospect theory can provide a richer and more holistic discussion of participants’ decision-making and strategy in crowd-based contests.

Results in this study also provide implications for firms that use crowd-based contests. First, firms need to be mindful of participants’ goals and cost considerations in contests. Due to relatively low chances of winning contests and their preference for cost minimization, participants’ may use strategies that are not aligned with firms’ interest. Such misalignments could limit the extent to which firms benefit from collective effort and wisdom through crowdsourcing. Therefore, firms should not solely use contest prizes to motivate participants, but also provide other meaningful incentives – unrelated to winning – to influence participants’ to submit better solutions. Second, firms need to be strategic when they interact with contest participants. The results show that although client-provided information does affect participants’ strategies and submissions, it could work in unexpected ways due to the nature of participants’ interest. As such, firms should recognize the objectives that matter to the types of individuals that take part in their contests. In the case of design contests, participants tend to be creativity-driven, and may be more concerned with submitting solutions that stand out from the competition than with winning and/or satisfying firms’ design preferences. Knowing how such participants might respond to client-provided signals would allow firms to better structure their communications (such as project briefs or feedback to submissions) with the crowd.

A limitation in this study is that I used a blind contest setting, where individual participants could not see other participants’ submissions during the contest. However, in practice, clients can also run open contests, where submissions are shown to participants during the contests. Such a contest feature allows participants to benchmark their submissions during the contest, which could affect their strategy and the resulting distinctiveness of their designs. Future research can examine how blind and open contests have different impacts on contestants’ strategy and submissions. Also, future studies can use qualitative data (e.g., interviews) to complement experimental and/or archival observations to better understand the cognitive processes by which participants interpret client-provided information in the contests.
References


