Understanding the effects of socialness and color complexity in listing images on crowdfunding behavior

Stuart John Barnes
Newcastle University Business School, Newcastle University, Newcastle upon Tyne, UK

Abstract

Purpose – Color psychology theory reveals that complex images with very varied palettes and many different colors are likely to be considered unattractive by individuals. Notwithstanding, web content containing social signals may be more attractive via the initiation of a social connection. This research investigates a predictive model blending variables from these theoretical perspectives to determine crowdfunding success.

Design/methodology/approach – The research is based on data from 176,614 Kickstarter projects. A number of machine learning and artificial intelligence techniques were employed to analyze the listing images for color complexity and the presence of people, while specific language features, including socialness, were measured in the listing text. Logistic regression was applied, controlling for several additional variables and predictive model was developed.

Findings – The findings supported the color complexity and socialness effects on crowdfunding success. The model achieves notable predictive value explaining 56.4% of variance. Listing images containing fewer colors and that have more similar colors are more likely to be crowdfunded successfully. Listings that convey greater socialness have a greater likelihood of being funded.

Originality/value – This investigation contributes a unique understanding of the effect of features of both socialness and color complexity on the success of crowdfunding ventures. A second contribution comes from the process and methods employed in the study, which provides a clear blueprint for the processing of large-scale analysis of soft information (images and text) in order to use them as variables in the scientific testing of theory.

Keywords Crowdfunding, Color complexity, Socialness, Big data, Machine learning

1. Introduction

Crowdfunding is fast becoming recognized as a formidable engine of employment and economic growth, with vast opportunities to mobilize resources via the internet. The worldwide market for crowdfunding was valued at $13.5bn in 2021, growing to $28.2bn by 2028 (Facts and Factors, 2022). It is perhaps not surprising then that crowdfunding has become a popular and important topic for information systems and business research more widely in recent years (Alhammad et al., 2022; Bargoni et al., 2022; Deng et al., 2022; Jiang et al., 2022; Mazzocchini and Lucarelli, 2023; Liu et al., 2023; Xiao et al., 2021). Considering the prominence of the crowdfunding market, understanding the elements that contribute to a greater likelihood of obtaining successful funding appears valuable. By recognizing such features, a project listing could be constructed in such a way as to have a greater probability of success. Thus, projects would be more likely to achieve the required amount of funding and therefore to be implemented in practice, reaping the associated economic returns.

Among many features of successful funding that have been investigated in the research, such as those related to the project, the sponsors, the creator, the platform, soft information, hard information, and other environmental factors, soft information factors have been found to be one promising driver of behaviors, and a focus of recent research studies (e.g. Hou et al.,...
Individual perceptions of a potential crowdfunding project from the features of a listing, including text and images, contribute to the development of pre-funding beliefs, in the same way that they do for a product or service (Oliver, 1980). Further research is needed to understand the psychological drivers of funding behavior, particularly using real, large-scale data from projects, rather than the small-scale surveys that are prone to many kinds of bias (Barnes, 2022). Color psychology theory implies that images containing disparate colors and complex palettes will be unattractive to individuals. Furthermore, web pages with embedded social cues may be more appealing through the development of a social connection with the viewer. This research investigates a predictive model blending variables from these theoretical perspectives to determine crowdfunding success. These features contribute directly to the backer’s decision regarding whether to fund or not. The research contributes to our understanding of the features of successful crowdfunding listings and helps provide directions for the establishment of crowdfunding listings that are more likely to be successfully funded.

Various scholars have examined the features underpinning crowdfunding success. Features investigated include textual sentiment and image color, but no study has examined the effect of the complexity of colors used in the principal images of crowdfunding listings (combined sets of colors) or the unifying idea of socialness within a listing via both text and image features. This investigation contributes a novel understanding of the effect of features of both socialness and color complexity on the success of crowdfunding ventures. The findings from this research support hypothesized relationships between both color complexity and socialness on successful project crowdfunding. To increase the odds of successful crowdfunding, project developers can emphasize social cues in text descriptions (e.g. social words) and listing images (e.g. human faces) for their project listings. Furthermore, listings with images containing fewer and more similar colors are more likely to be crowdfunded successfully. A second contribution comes from the process and methods employed in the study, which provides a clear blueprint for the process of large-scale analysis of soft information (images and text) in order to use them as variables in the scientific testing of theory. The research process uses data from Kickstarter to measure the socialness of listings via face detection in listing images and analysis of social words in listing text. Color complexity is assessed by extracting the main colors in a listing image using k-means cluster analysis and examining palette diversity and the perceptual distance between the main colors. Various control variables are included for the image, text, and listing. The data is combined and analyzed in a final model using logistic regression to predict funding success.

The structure of the paper is as follows. The subsequent section explores the literature foundation, including the justification of hypothesis and establishment of a research model. The third section delineates the methodology used in the study, focusing on a detailed explanation of the elements of the research process. The fourth section tests the research model and evaluates its value from a predictive perspective. The final section reflects on the research limitations and implications for practice and further research.

2. Literature foundation and hypotheses

2.1 Crowdfunding success in the literature

Research examining features underlying the success in funding projects on crowdfunding platforms has growing steadily in the past two decades. Table 1 summarizes some recent research examining crowdfunding success.

The overwhelmingly prevalent dependent variable employed in previous empirical studies is successful funding of projects, whilst the dominant analytical methods are linear and logistic regression (Deng et al., 2022). Numerous recent studies have taken a more holistic view of the corpus of previous research into crowdfunding success factors, including via
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of study</th>
<th>Success factors examined</th>
</tr>
</thead>
</table>
| Koch and Siering (2015)  | Empirical study of projects on Kickstarter (logistic regression with n = 762) | • Project-specific media factors (depth of project description, graphical accompaniment and provision of video material)  
• Other project-specific factors (availability of project updates and size of pledging goal)  
• Funder-specific aspects (project experience and funding reciprocity)  
• Content of project description (length, readability and tone)  
• Trustworthy cues (past experience and past expertise)  
• Other variables (goal, duration, connected to Facebook, Facebook friends, number of images, number of videos, number of pledge levels, year launched and category of project) |
| Zhou et al. (2018)       | Empirical study of project description factors for 151,752 projects via logistic regression | • Commitment, communication, funding intention, perceived benefit, perceived innovation, perceived risk, product involvement, shared value and trust  
• Pledge conditions (funding goal and funding period)  
• Information disclosure (no. of words, no. of words squared, no. of pictures, no. of pictures squared and provision of title video)  
• Risk disclosure (no. of risk words and no. of risk words squared)  
• Emotional appeal (text polarity)  
• Project experience (previous crowdfunding campaigns and number of previous successful projects)  
• User popularity (log of Facebook friends)  
• Hard information (reward execution, initiator experience in creating projects, number of failed projects, number of projects supported, number of backers and number of comments)  
• Soft information (topic ratio, positive sentiment and negative sentiment)  
• Controls (target funds, reward level, minimum pledge set, maximum pledge set, category, financing model and region) |
| Zhao et al. (2017)       | Empirical study of 204 projects via partial least squares path modeling        | • Commitment, communication, funding intention, perceived benefit, perceived innovation, perceived risk, product involvement, shared value and trust  
• Pledge conditions (funding goal and funding period)  
• Information disclosure (no. of words, no. of words squared, no. of pictures, no. of pictures squared and provision of title video)  
• Risk disclosure (no. of risk words and no. of risk words squared)  
• Emotional appeal (text polarity)  
• Project experience (previous crowdfunding campaigns and number of previous successful projects)  
• User popularity (log of Facebook friends)  
• Hard information (reward execution, initiator experience in creating projects, number of failed projects, number of projects supported, number of backers and number of comments)  
• Soft information (topic ratio, positive sentiment and negative sentiment)  
• Controls (target funds, reward level, minimum pledge set, maximum pledge set, category, financing model and region) |
| Koch and Siering (2019)  | Empirical study of projects on Kickstarter (logit regression with n = 32,083) | • Project-specific media factors (depth of project description, graphical accompaniment and provision of video material)  
• Other project-specific factors (availability of project updates and size of pledging goal)  
• Funder-specific aspects (project experience and funding reciprocity)  
• Content of project description (length, readability and tone)  
• Trustworthy cues (past experience and past expertise)  
• Other variables (goal, duration, connected to Facebook, Facebook friends, number of images, number of videos, number of pledge levels, year launched and category of project) |
| Jiang et al. (2020)      | Empirical study of 916 projects in China via linear regression                 | • Hard information (reward execution, initiator experience in creating projects, number of failed projects, number of projects supported, number of backers and number of comments)  
• Soft information (topic ratio, positive sentiment and negative sentiment)  
• Controls (target funds, reward level, minimum pledge set, maximum pledge set, category, financing model and region)  
(continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of study</th>
<th>Success factors examined</th>
</tr>
</thead>
</table>
| Schneor and Vik (2020)   | Systematic literature review of 88 papers          | - Fundraiser-related factors, such as location, cultural dimensions, sector, IPR ownership, credibility and reputation, team/staff size, team openness, ethnic diversity in team, gender, social media presence, network size and social capital  
- Platform-related factors (Crowdfunding industry associations, brand indicators, crowdfunding model, flexibility and openness, reputation, following, success rates, platform age at campaign launch and recency of campaign)  
- Campaign-related factors, such as tax incentives, regulation, political regime, competition intensity on platform, writing style, video indicators, image indicators, number of images, text length and features, number of backers, reward type, investment indicators, social media engagement by crowd and campaign uncertainty indicators  
- Concept-level features (Purpose: social finance, personal finance or business finance, creativity and innovativeness, complexity, project maturity, quality, target market, scalability, perceived attractiveness and perceived relatability)  
- Funder-related features (social trust, online trust, gender, same gender as creator, subjective evaluation admission, involvement in project beyond dollars and low effort) |
| Deng et al. (2022)       | Systematic literature review of 94 papers          | - Project-related factors: project characteristics, such as goal, early funds, early backers, total backers, pledge/backer ratio, duration, reward, team size and category  
- Project-related factors: soft information (text, visual, social network, updates, comment, staff pick, shares, recommended, likes, media coverage, signals, followers and FAQs)  
- Creator-related factors, such as preparedness, innovativeness, user entrepreneurship, experience, education, patent ownership, Facebook friends, culture, geography, race, firm age and prior funding  
- Backer-related factors (funder positive affective reactions, motive, experience, platform tenure and geography)  
- Platform-related factors (competition, platform type and platform age)                                                                                                        |
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of study</th>
<th>Success factors examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geiger and Moore (2022)</td>
<td>Meta-analytic structural equation modeling using 112 samples</td>
<td>• Text (number of words or characters used in a campaign)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Images (number of images or photos)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Videos (number of videos)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Positive tone (number positive emotion words)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Gender</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Number of backers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Funding goal</td>
</tr>
<tr>
<td>Jáki et al. (2022)</td>
<td>Literature review of 53 papers</td>
<td>• Campaign attributes (financial planning, pledges, project scope, product, rewards, platform features and timing)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Characteristics of participants (initiators intellectual capital, initiator's crowdfunding experience, initiator's personality, other characteristics of the initiator and backer's characteristics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Communication (initiator-backer interactivity, content, semantics and media elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Network (internal network and external network)</td>
</tr>
<tr>
<td>Li et al. (2022)</td>
<td>Empirical study of 2,877 projects in the cultural and creative sector via probit regression</td>
<td>• Peer signal (peer review valence and information entropy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Project signal (project cultural background)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Initiator signal (face information disclosure)</td>
</tr>
<tr>
<td>Liu et al. (2022)</td>
<td>Meta-analysis of 53 papers</td>
<td>• Project-related factors (goal, duration, description, video, image and quality)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fundraiser-related factors (age, gender, experience, location, friends, team and ethnicity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Platform-related factors (likes, comments, updates, staff pick, shares, followers and available links)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Investor-related factors (age, gender, experience, investors, community and reward)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of study</th>
<th>Success factors examined</th>
</tr>
</thead>
</table>
| Hou et al. (2023)     | Empirical study of 840 charity-giving projects on Kickstarter using regression. Experiment with 177 respondents from MTurk | - Image emotions (amusement, awe, anger, contentment, disgust, excitement, fear and sadness)  
- Competing project image emotions (amusement, awe, anger, contentment, disgust, excitement, fear and sadness)  
- Image content (animal in image, human in image, human face in image and text in image)  
- Composition (diagonal dominance, symmetry, visual balance of color and rule of thirds)  
- Main element-background relationship (size difference, color difference and texture difference)  
- Emotion in text description (anxiety, sadness and anger)  
- Project characteristics (pre-set goal, project popularity, length of text description, number of images, number of videos and campaign duration)  
- Color (warm hue, saturation, brightness and contrast of brightness)  
- Crowdfunding features (project creator, social information, project description and rewards)  
- Mechanisms (ties strength, credibility of creator, perceived project quality, collect tangible reward, collect intangible reward, emotional reaction and identify with community) |
| van Teunenbroek et al. (2023) | Thematic literature review of 198 papers in donation-based crowdfunding |                                                                                                                                                             |

Source(s): The author's own creation/work
meta-analysis (Geiger and Moore, 2022; Li et al., 2022) and literature reviews (Deng et al., 2022; van Tuenenbroek et al., 2023; Jáki et al., 2022; Shneor and Vik, 2020). Success factors are many and varied but can generally be broken down into those related to characteristics of the concept, project and project campaign, factors focusing on the creator/initiator/fundraiser, project listing-related factors (including soft and hard information such as metrics, text, images, and videos), factors related to backers/supporters/donors/investors/funders, and factors related to the platform/network.

Within the set of soft information features explored in the crowdfunding studies examining determinants of funding success, textual and visual elements have tended to be most prevalent, with a sizeable body of studies supporting the effect of various dimensions of these features. Textual features analyzed are word count, with varying support (e.g. Geiger and Moore, 2022; Koch and Siering, 2019; Wang et al., 2020; Zhou et al., 2018), linguistic style (e.g. Wang et al., 2021), quality of written text, readability of and errors within text (e.g. Scheaf et al., 2018), sentiment expressed (e.g. Geiger and Moore, 2022; Jiang et al., 2020; Tafesse, 2021), and particular terms assessed through dictionary methods, including community and risk. Research into visual features have captured image quality, total number and accessibility of video and images (e.g. Chen et al., 2016; Koch and Siering, 2019; Geiger and Moore, 2022; Shneor and Vik, 2020). A handful of studies go further than this to investigate a few features of color and aspects such as emotion and humans in images and text (e.g. Guo et al., 2022; Hou et al., 2019, 2023; Li et al., 2022; Yazdani et al., 2022). Notwithstanding, no research has been found investigating the complexity of colors (combined colors) or a unifying theoretical construct that captures the social nature (e.g. social cues) of a crowdfunding listing in a meaningful way.

2.2 Socialness of websites
This study revolves around the concept of the degree to which a crowdfunding website exhibits socialness, perceived via different social cues in varied media. Social presence theory examines degree to which a communication medium may carry social cues (Gefen and Straub, 2003; Short et al., 1976). Notwithstanding, the salient literature has advanced to include the notion of website media (and computers more generally) as social artefacts in their own right, such that individuals mindlessly apply expectations and social rules in a similar way as to other human actors (Kumar and Benbasat, 2002; Moon, 2000; Nass et al., 1995).

Studies have demonstrated that user beliefs regarding information technology artefacts’ behaviors can be created via interaction, such as social beliefs and beliefs regarding actors’ relationships with IT artefacts, which is consistent with theories of interpersonal interaction (Al-Natour and Benbasat, 2009). Previous research has found that the degree of socialness of websites is strongly associated with satisfaction, a significant determinant of behavioral intention (Barnes and Vidgen, 2014). Even though a website might be considered as an abstract social artifact, it is capable of generating a sensation of social contact. These sensations include “a psychological connection with the user, who perceives the Web site as warm, personal, sociable, thus creating a feeling of human contact” (Cyr and Head, 2008) and “the perception of social presence can still be created despite the lack of actual human contact” (Gefen and Straub, 2004). Thus, numerous studies have subsequently been developed investigating website socialness and the related effects on user behavior (e.g. Cyr et al., 2007; Hassanein and Head, 2005-2006; 2007; Kumar and Benbasat, 2002; Patilollu, 2023; Zhou and Jia, 2018). The concept has been extensively examined in the context of social media websites (Dabiran et al., 2022; Kumar and Singh, 2022; Zhong et al., 2021). Dabiran et al. (2022) examine the impact of four anthropomorphic elements in social media – appearance, moral virtue, cognitive experience, and conscious emotionality – and their effects on followers’ purchase intention. Zhong et al. (2021) examine the impact of parasocial relationships on
brand-consumer relationships in social media. A parasocial relationship refers to a connection – an interest, feelings of familiarity or closeness – between an individual and another person or entity. Using a survey of 209 respondents and structural equation modeling, Zhong et al. (2021) test a research model and find significant relationships between utilitarian and hedonic benefits and parasocial relationship, and between parasocial relationship and brand engagement and brand loyalty.

A recent focus of parasocial research into websites has been a focus on the inclusion of artificial intelligence-powered chatbots and voice assistants on websites (Beattie et al., 2020; De Cicco et al., 2020; Lee and Park, 2022; Whang and Im, 2020). For example, Lee and Park (2022) examine how interactions with shopping chatbots influences communication quality, satisfaction, and continuance intentions for 184 women in Korea using partial least squares path modeling. They find a strong relationship between a consumer’s parasocial relationship with the chatbot and continuance intention via satisfaction and communication quality. Using two experiments with 84 and 418 respondents, Whang and Im (2020) find that consumers tend to perceive voice assistants as pseudo-human agents separate from a service provider, creating a more positive perception and evaluation of websites.

Focusing on a retailer’s website in China, Zhou and Jia (2018) examine the role of parasocial interaction and psychological distance in mediating the relationship between website quality and relationship quality. A survey of 223 respondents and covariance structural equation modeling found that the parasocial interaction fully mediated the relationship between website quality and relationship quality. Zhou et al. (2021a) found that animated spokes-characters in logos can generate parasocial interaction relationships via likeability, expertise, and congruence, which can positively drive ad effects (ad attitude, brand attitude, and purchase intention), with some degree of mediation. In the context of social shopping websites, Zheng et al. (2020) examined the impact of three types of technology attraction, social, physical, and task on parasocial interaction and social commerce intention. Using a survey of 248 website users and partial least squares structural equation modeling they find that social attraction and task attraction drive parasocial interaction, which in turn has a significant relationship with social commerce intention. In the tourism context, Chen et al. (2022) examined the notion of parasocial interaction, self-congruity, and perceived authenticity with influential content from online travel industry websites in China, as a driver of source credibility, destination and website trust, and ultimately behavioral intention. A survey of 420 students was conducted and analyzed using covariance structural equation modeling, finding parasocial interaction as a significant driver of perceived trustworthiness and perceived expertise.

Overall, the previous research demonstrates that individuals can exhibit strong parasocial relationships towards technology artefacts such as websites. Moreover, the body of research suggests that the socialness exhibited by websites can influence individual attitudes and behaviors. In general, socialness is typically examined in the literature using social psychological constructs measured via survey and tested using structural equation modeling or regression. Data analytics of text, images, and video using big data sets can help ameliorate various aspects of bias that may be prevalent in surveys (see Barnes et al., 2020 for a summary), including sampling error (Dillman et al., 2014; Singleton et al., 2009), social desirability bias (De Vaus, 1996), recall bias (De Vaus, 1996), common method bias (Podsakoff et al., 2003), inattention bias (Brosnan et al., 2019), and measurement error (Dillman et al., 2014).

There is still of lack of research employing big data analytics to investigate socialness, with a few rare studies measuring the presence of a face in the crowdfunding context, such as Hou et al. (2023) and Li et al. (2022), and no study examine the presence of socialness in text for crowdfunding listing images. Studies examining the presence of faces lack a unifying theoretical foundation for socialness.
In this study, we consider a listing on a crowdfunding website as a social entity and social artefact that conveys a sense of socialness to potential crowdfunding backers that may influence their behavior. Individuals consider computers and websites as social artefacts and mindlessly apply expectations and social rules in a manner similar to that exhibited towards other humans (Kumar and Benbasat, 2002; Moon, 2000; Nass et al., 1995). Websites may develop a psychological connection and social presence with viewers through a feeling of human contact (Cyr and Head, 2008; Gefen and Straub, 2004).

Individuals may develop parasocial relationships with technology such as websites, chatbots, and other technologies with embedded social cues (Beattie et al., 2020; Jiang et al., 2020; Whang and Im, 2020; Zheng et al., 2020). The level of socialness exhibited by websites is significantly associated with satisfaction, and subsequently behavioral intention (Barnes and Vidgen, 2014). Given that crowdfunding listings can be considered as websites, and that they have embedded social cues, for example through the images and text used in a listing, we would expect them to exhibit a degree of socialness. Images may include human faces and listing descriptions may be written in a way that exhibits greater socialness. Great socialness in crowdfunding listings is likely to drive more positive behavior—in this study a greater likelihood of funding from backers. Consistent with this, it is hypothesized that:

**H1a.** The proportion of social text listing descriptions has a positive relationship with funding success.

**H1b.** The presence of a face in listing images has a positive relationship with funding success.

**H1c.** The presence of multiple faces in listing images has a positive relationship with funding success.

2.3 Complexity of colors

Although color has been considered in previous studies examining crowdfunding success (e.g. Guo et al., 2022; Hou et al., 2019, 2023; Yazdani et al., 2022), the variables examined are limited with respect to color complexity and sample sizes. Hou et al. (2023) examined the impact of warm hue, saturation, brightness, and contrast of brightness in their analysis of factors determining charity crowdfunding on Kickstarter for 840 projects. Their underpinning theory is that of basic stimulus-organism-response (S-O-R). Hou et al. also examine visual balance of color and color difference between the main element and the background, based on the work of Zhang et al. (2021). Euclidian color distance (a very limited measure—as we shall see below) is calculated between the average color of the body and the average color of the background. Visual balance of color is measured based on the additive inverse of the mean of Euclidian distance between mirrored pixel pairs on the vertical line. Although they do not examine the impact of these measures on project success, Hou et al. (2023) find a relationship with some emotions. Color balance has significant (1% level) relationships with awe (positive) and with disgust (negative). Color difference has a significant positive relationship with anger (1% level).

This research contributes knowledge towards a gap in the literature regarding understanding the effect of color complexity on individual behavior, focusing on successful funding of crowdfunding ventures. This topic has had minimal attention in the literature, and only a few studies exist, with two focusing on the product context (Deng et al., 2010; Van Kerckhove and De Bock, 2014) and a single study focusing on accommodation (Barnes, 2022). Listing images play an important role in the development of individual impressions regarding products and services offered on websites (Barnes, 2022). Such impressions occur very rapidly (Barnes and Kirshner, 2021). Understanding the effect of color complexity on individual behavior is important because it enables more effective image
creation and curation to maximize positive image impression on potential backers and thus the likelihood of project funding.

Gestalt psychology principles of unity and similarity suggest that elements of “good” images tend to look similar and fit together through visual connection as opposed to mere chance (Koffka, 1935; Kumar and Garg, 2010). Graphical designers typically try to create this aesthetic via diligent matching of colors that appear similar in a carefully curated palette (Kim and Suk, 2018). A key explanation for an individual’s psychological preference is that the brain tries to lessen the utilization of attentional resources to improve processing efficiency (Hekkert, 2006). Consequently, visual representations that employ principles of unification and similarity, via the inclusion of design elements that are more alike and fewer in number, including color, provide a neuropsychological advantage (Kumar and Garg, 2010).

Research in the area of services has also identified that highly complex visuals can lessen attractiveness (Orth and Wirtz, 2014). Consumers prefer sets of colors matching closely in hue and chroma (Deng et al., 2010), and that are perceived as being in harmony rather than disharmony (Bell et al., 1991). Moreover, research finds that consumers are more willing to pay when considering product packages consisting of similar as opposed to contrasting colors (Van Kerckhove and De Bock, 2014).

Empirical evidence indicates that consumers prefer smaller color palettes. Research in the self-design context has offered scientific support for the small palette principle (Deng et al., 2010). In combination, the principles of similarity and unity suggest that good images consist of a lower number of colors, as colors closer in chroma and hue are preferred by individuals in the processing of a whole image (Wei et al., 2014; Schloss and Palmer, 2011).

Consumer expectations and beliefs may be affected by observing the features of a product or service (Oliver, 1980). Within an online environment, expectations can be developed via elements from a great variety of sources, such as images, textual descriptions, ratings, previous experience, advertising, social media, brand image, and reviews, each of which has the potential to have an effect on behavior. In a crowdfunding context, one would anticipate that backer expectations are developed partly via the processing of color features from a project listing image, and that consequently this effects a backer’s decision-making with respect to the contribution of funds.

In consideration of the above, it is hypothesized that the visual complexity of listing images for a crowdfunding project, manifested via a greater color palette and lower color coherence, will have a negative effect on successful funding of the project. Thus, it is posited that:

\[ H2a. \] Palette simplicity in listing images has a positive relationship with funding success.

\[ H2b. \] Lack of color coherence in listing images has a negative relationship with funding success.

The research model tested is shown in Figure 1. The socialness aspect of Figure 1 is captured via three variables testing the effect of the socialness of textual and visual features on successful crowdfunding (H1a, H1b and H1c) and two variables testing the effect of the complexity of colors in listing images on successful crowdfunding (H2a and H2b). The research model includes controls for various additional variables, including the word count of the descriptive listing text, the proportion of informal text, the number of project backers, overall image quality, and overall luminance of the image used in a listing.

### 3. Methodology

The research process and its constituent sequential steps are shown in Figure 2. Each of these will now be explained.
3.1 Data preparation

The study data set was based on projects seeking funding via Kickstarter, the crowdfunding platform with the greatest number of projects. All study data were available publicly from Web Robots (http://webrobots.io/kickstarter-datasets/), containing an archive of data for all Kickstarter projects since March 2016. The data collected included a description for each listing, a web link for each listing image, as well as additional data on each project. The data set was assembled in May 2021. From a total of 224,518 listing images, after cleaning the data for duplicates, 194,618 images were left.
3.2 Face detection
To detect faces, the Viola–Jones algorithm was employed to analyze the listing images in MATLAB. This algorithm has been used in similar studies (e.g. Lu et al., 2016). The analysis involves a cascade classification procedure utilizing local binary patterns to ensure robustness against illumination disparity (Ojala et al., 2002), determined by a pre-assessment of listing images. Face detection was used to create two variables: firstly, the discovery of at least a single face; and secondly, the observation of more than one face.

3.3 Image luminance and quality assessment
A number of control variables based on image features were used in the study. Image quality typically refers to how clear or sharp the graphic is perceived. In this study the Perception-based Image Quality Evaluator (PIQE) algorithm (Venkatanath et al., 2015) was used to calculate image quality in MATLAB. PIQE is based on a no-reference image quality metric. No-reference quality evaluations generally outperform full reference assessments in their similarity to human perception. PIQE metrics are scored from 0 to 100, where 0 is the highest quality image and 100 is the poorest quality image. Luminance refers to the magnitude of light emitted from an image; this metric was calculated via the colordistance package (Weller, 2022) and schemr (Morrison, 2022).

3.4 K-means cluster analysis for color extraction
The main colors from each image were extracted via the following process. Firstly, an assessment was required to determine the appropriate number of colors in each listing image. The number of colors to be extracted, the choice of k in the analysis, was informed via a silhouette score calculation for each image in the data set. The silhouette score of datum point i is given as:

$$\varphi_i = \frac{b_i - a_i}{\max\{a_i, b_i\}} \text{ if } |\gamma_i| > 1$$

where $-1 \leq \varphi_i \leq 1$, each datum point i is part of a cluster, $i \in \gamma_i$, $b_i$ is the shortest distance between i and every other datum point in other clusters, $a_i$ refers to the mean distance between i and all data points in its own cluster, $\gamma_i$, and $\varphi_i = 0$ if $|\gamma_i| = 1$. Secondly, applying (1), the silhouette score is the maximum mean value $\varphi_i$ over all datum points for image v such that:

$$\text{Silhouette}_v = \max(\zeta_k)$$

where $\zeta_k$ refers to the mean of $\varphi_i$ for all image datum points across k clusters (Kaufman and Rousseeuw, 1990).

For all images in the data set, the mean Silhouette$_v$ was calculated as 0.732 (SD = 0.128). The mean value of k was 3.429 (SD = 2.453). Based on this assessment, a value of k = 10 would be able to capture 99.63% of all color clusters. Subsequently, 10 colors were extracted from each image in the data set via k-means cluster analysis. The method determines color clusters by the minimization of the sum of all distances between pixels in a color space, from which it is possible to calculate cluster centers. The colorspace package (Zeileis et al., 2020) was used for the analysis.

3.5 Determining main color types
A mapping process was used to determine the principal type of each color (Barnes, 2022). First, each of f extracted colors from every listing image was mapped via Euclidean distance
from its original RGB color space to its equivalent $R$ color class. Second, each one of $f$ colors was mapped from its $R$ color class to one of 12 main colors (black, blue, brown, cyan, green, grey, orange, pink, purple, red, red, white, and yellow) via color mappings (Rapid Tables, 2021). Third, the totals of colors for each main color group were calculated for each image and converted to binary variables, to indicate the presence of the main colors.

### 3.6 Assessing palette simplicity and color coherence

A popular color space used in scientific study is CIELAB; a device-independent color space developed by the Commission Internationale de L’Eclairage (International Commission on Illumination) in 1976. CIELAB represents color by means of perceived lightness ($L^*$) and four unique colors in our vision – green and blue, red and yellow – whereby the $a^*$ axis represents opposite green to red colors and the $b^*$ axis represents yellow to blue juxtaposing colors. The 1976 CIELAB $\Delta E$ measurement has been found to have approximately 75% agreement with human vision. A more advanced measurement was created by CIELAB in 1994, $\Delta E^{*}_{94}$. This second metric has around 95% agreement with human vision, a considerable improvement, and was selected for use in this study (Barnes, 2022). Existing studies examining color distance have employed Euclidian distance, which is a very basic and inferior measure of color distance, and not to be recommended for scientific research. It does a very poor job of measuring in terms of agreement with the human eye (Barnes, 2022).

When comparing two colors in our calculations, it is important to account for the proportional area of an image covered by each of the colors. Thus, the $\Delta E^{*}_{94}$ calculations were weighted according to the percentage of an image covered the two colors being compared, $w\Delta E^{*}_{94}$. This follows best practice in prior research (e.g. Barnes, 2022; Li et al., 2020). The final metric for color coherence is calculated as the mean distance between each pair of $f$ colors extracted using k-means from an image. Since $k = 10$, this becomes the mean for the comparison of 45 color pairs, given by:

$$w\Delta E^{*}_{94} = \frac{\sum_{q=1}^{f(f-1)/2} w\Delta E^{*}_{94,ij}}{f(f-1)/2} \quad (3)$$

where $q$ is the unique color comparison, 1 to 45, between each color i and j. In the following analysis, this variable is referred to as $WE94$.

Color palette simplicity was determined via the Normalized Herfindahl-Hirschman Index (NHHI). This measure was first created to assess industry concentration (Hirschman, 1964), and is analogous to the Simpson Index measuring diversity in the discipline of ecology (Simpson, 1949) and the Blau index in the discipline of sociology (Blau, 1977). NHHI (and its equivalents) is widely used in science and has recently been applied to color science (Barnes, 2022). NHHI in this study measures the number of $f$ colors within each of $z$ main color groups. It is calculated as:

$$NHHI = \frac{\left(\sum_{m=1}^{z} \rho_m z - 1 \right) / z}{1 - 1/z} \quad (4)$$

where $m \in \{1, 2, 3, \ldots z\}$ and $\rho$ relates the share of $f$ main colors in color group $m$.

Figure 3 provides two examples of the measurement of image-related variables for two images in a similar way to Zhou et al. (2021b). These images differ in terms of several of the variables measured, including face detection, image quality, palette simplicity, and luminance.

Effects of socialness and color complexity
3.7 Measuring textual variables
To assess the degree of informality and socialness in the text of project listings, predetermined dictionaries from the LIWC2015 software package (Pennebaker et al., 2015a, b) were employed. LIWC2015 has been developed over more than 30 years and utilized in hundreds of scientific studies in a broad range of disciplines. LIWC’s dictionaries have proven to be robust for analysis of text. The English LIWC2015 contains more than 6,500 words across various categories. The dictionaries include 756 social words (e.g. buddy, her, mate, they, talk, daughter, and dad) and 380 words related to informal language (assent, netspeak, non-fluencies. Swear words, and fillers) (Pennebaker et al., 2015b). The description of each project listing was analyzed via LIWC2015, resulting in a percentage score for informal and social words used. At the same time, LIWC2015 calculated word counts for each project description.

3.8 Combining text, image and additional data
Data measured from project listing images (i.e. WE94, NHHi, presence of a face/multiple faces, luminance, and image quality) were merged with calculated variables from listing text (i.e. word count, social text, and informal language) and additional data from the projects (i.e. successfully funded: 0 = no, 1 = yes; and the number of backers). Only projects with complete data were included in the analysis (n = 176,614 projects).

3.9 Hypothesis testing and predictive assessment
The dependent variable in the study is binary and therefore binary logistic regression was employed for model testing. An advantage of logistic regression is that it permits the calculation of the reduced or increased odds for crowdfunding and the evaluation of the predictive value for our classification model. All of the variables in the research model were tested in a single regression; the dependent variable was funding success. The predictive value of the research model was evaluated by means of pseudo-$R^2$ metrics, a confusion matrix, and ROC (Receiver Operating Characteristic) analysis utilizing the probabilities calculated from the model analysis.
4. Results and discussion

4.1 Sample descriptive statistics

Descriptive statistics on the sample are summarized in Table 2. There was a complete data set for 176,614 projects seeking Kickstarter funding. Approximately three-fifths of projects were funded. Some 30.7% of the listing images contained a face, while 10.5% included more than one face. The mean word count of the listing text was 18 words, with a mean of 8.3% social text and 54.8% informal text. The number of backers varied widely, with a mean of 155 and a standard deviation of 968 persons.

4.2 Testing the research model and discussion of results

The results of the binary logistic regression analysis are summarized in Table 3. With respect to the socialness variables included in the model, the regression found significant positive relationships for all three with successful project funding. The presence of more than one face in a crowdfunding project listing image was the most significant of the variables ($\beta = 0.109$, SE = 0.025, $p < 0.001$); the presence of more than one face led to a rise in odds by a factor of

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta$</th>
<th>S.E.</th>
<th>Wald</th>
<th>$p$-value</th>
<th>Exp($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.734</td>
<td>0.041</td>
<td>317.599</td>
<td>&lt;0.001</td>
<td>0.480</td>
</tr>
<tr>
<td><strong>Color complexity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WE94</td>
<td>-0.205</td>
<td>0.063</td>
<td>10.415</td>
<td>0.001</td>
<td>0.815</td>
</tr>
<tr>
<td>NHHI</td>
<td>-0.107</td>
<td>0.031</td>
<td>11.759</td>
<td>0.001</td>
<td>0.888</td>
</tr>
<tr>
<td><strong>Parasocialness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social text</td>
<td>0.002</td>
<td>0.001</td>
<td>7.367</td>
<td>0.007</td>
<td>1.002</td>
</tr>
<tr>
<td>Presence of a Face</td>
<td>0.057</td>
<td>0.017</td>
<td>11.521</td>
<td>0.001</td>
<td>1.058</td>
</tr>
<tr>
<td>Multiple faces</td>
<td>0.109</td>
<td>0.025</td>
<td>18.953</td>
<td>&lt;0.001</td>
<td>1.115</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word count</td>
<td>-0.015</td>
<td>0.001</td>
<td>160.862</td>
<td>&lt;0.001</td>
<td>0.999</td>
</tr>
<tr>
<td>Informal text</td>
<td>-0.016</td>
<td>0.003</td>
<td>27.131</td>
<td>&lt;0.001</td>
<td>0.995</td>
</tr>
<tr>
<td>Number of backers</td>
<td>0.053</td>
<td>&lt;0.001</td>
<td>29042.871</td>
<td>&lt;0.001</td>
<td>1.064</td>
</tr>
<tr>
<td>PIQE</td>
<td>-0.0004</td>
<td>&lt;0.001</td>
<td>108.304</td>
<td>&lt;0.001</td>
<td>0.996</td>
</tr>
<tr>
<td>Luminance</td>
<td>-0.002</td>
<td>&lt;0.001</td>
<td>44.827</td>
<td>&lt;0.001</td>
<td>0.998</td>
</tr>
</tbody>
</table>

**Source(s):** The author's own creation/work

Table 2. Descriptive statistics for the sample

Table 3. Testing of the research model
The presence of a single face in a crowdfunding listing image was also significant in the model ($\beta = 0.057$, SE = 0.017, $p = 0.001$), leading to a change (increase) in the likelihood of being funded by a factor of 1.058 (5.8% above parity). This provided support for H1b. The presence of social text in a project description had a significant positive relationship with success in project funding ($\beta = 0.002$, SE = 0.001, $p = 0.007$), but the increase in odds is small – by a factor of 1.002. Notwithstanding, the finding supports H1a.

The results of testing for the presence of social variables in the research model supports the notion that the crowdfunding listings can exhibit socialness that exerts an impact on the viewers, providing impetus for funding of projects. The literature suggests that parasocial relationships and interactions can be developed fruitfully with websites and social media (Chen et al., 2022; Zheng et al., 2020; Zhou and Jia, 2018; Zhou et al., 2021a). This research demonstrates that socialness theory holds in the context of crowdfunding platforms and the results provide a potential lever for influencing project success by careful design of listing content.

Turning to the block of variables measuring the effect of color complexity on the probability of successful crowdfunding, we see that the two variables measured have negative relationships and are significant. Lack of color coherence demonstrates the greatest negative relationship with successful funding ($\beta = -0.205$, SE = 0.063, $p = 0.001$), whereby the odds change (fall) by a factor of 0.815 for every unit increase in WE94 (18.5% below parity). This finding supports H2b. Palette simplicity (evaluated via the Normalized Herfindahl-Hirschman Index), has a significant negative effect on successful crowdfunding ($\beta = -0.107$, SE = 0.031, $p = 0.001$), with the odds of funding changing (falling) by a factor of 0.898 (10.2% below parity). This finding supports H2a.

The findings of the research support the notion that individuals prefer images that display fewer colors and that are more similar (have less complexity in terms of color distance between the main colors). This is underpinned by Gestalt psychology (principles of similarity and unification), the brain’s preference for processing efficiency (Hekkert, 2006), and generally implemented as good practice in the graphic design context. Although no similar research exists, in service environment, one study found that high visual complexity can reduce attractiveness (Orth and Wirtz, 2014). Other research in the product context has found that consumers tend to favor colors that match closely in chroma and hue (Deng et al., 2010) and that are considered in harmony rather than in disharmony (Bell et al., 1991). Indeed, there is some tentative research that suggests that consumers are more willing to pay for product packages where colors are similar rather than contrasting (Van Kerckhove and De Bock, 2014).

No research has considered color complexity and project success in the crowdfunding context. Hou et al. (2023) measure the visual balance of color on the vertical line of images and color difference between the main element and the background in charity crowdfunding projects. They find a relationship between these variables and some emotions (awe and anger; anger respectively, at the 1% level). Hou et al. (2023) do not examine the relationship between these measures and project success. Barnes (2022) found a relationship between the color complexity of listing images and visitor patronage via Airbnb data, implementing both palette simplicity via NHH and advanced recent CIELAB measures. Zhang et al. (2021) implement the same measures as Hou et al. (2023) to examine the relationship between vertical color balance and figure-ground color difference using very basic Euclidian distance measures. Vertical color balance had a very strong relationship with property demand ($p < 0.001$), while figure-ground color difference was also found to have a significant effect ($p < 0.05$).

Each control variable had a significant effect on successful crowdfunding. The largest effect was that of the number of backers ($\beta = 0.053$, SE < 0.001, $p < 0.001$), whereby the odds
of funding changed by a factor of 1.054 (5.4% above parity). The additional control variables had much smaller effects on funding success, changing the odds of successful crowdfunding by factors of 0.998 (0.2% below parity) and 0.985 (1.5% below parity).

The $\chi^2$ Omnibus Test of Model Coefficients was found to be significant ($p < 0.001$), implying that the research model is a good fit. The pseudo-$R^2$ values are sizable, with the Nagelkerke $R^2$ metric at 56.4% and the Cox and Snell $R^2$ metric at 41.8%. Nagelkerke is generally preferred in logistic regression analysis since it is more similar to $R^2$ in OLS regression, varying from 0 to 1, through a rescaling of the Cox and Snell $R^2$ metric (Nagelkerke, 1991). This implies that the research model, incorporating the concepts of color complexity and socialness, explains more than half of success for crowdfunding projects.

### 4.3 Predictive model assessment

Evaluation of the predictive value of the research model was initially examined via a confusion matrix, summarized in Table 4. The confusion matrix shows that the logistic model predicts successfully almost 90% of the total of projects that did not receive funding (89.8%), though the figure is less for the estimate of projects that were crowdfunded successfully at 81.7%. Overall, model accuracy is 85%.

For the ROC analysis, 104,447 of the projects processed were positive, whilst 72,167 were negative. Figure 4 shows that the predictive value for the logistic model is substantial, and the

<table>
<thead>
<tr>
<th>Predicted State</th>
<th>Unfunded</th>
<th>Funded</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual state</td>
<td>Unfunded</td>
<td>64,839</td>
<td>7,328</td>
</tr>
<tr>
<td>Overall percentage</td>
<td>Funded</td>
<td>19,137</td>
<td>85,310</td>
</tr>
</tbody>
</table>

Table 4. Confusion matrix for logistic regression model

Source(s): The author’s own creation/work

![ROC Curve](image)

Source(s): The author’s own creation/work
area under the curve (AUC) is found to be 0.939 (SE = 0.001, p < 0.001). The random model is indicated by the red diagonal line. The research model is represented by the blue curve. Sensitivity is measured via the standard formula: total true positives divided by true positives plus false negatives; specificity is calculated as the total of true negatives divided by true negatives plus false positives. The research implies that although the model is simplistic, it has predictive value in understanding the elements determining successful crowdfunding: the results suggest that color complexity and socialness, in addition to the other features measured, are at least in part important in developing projects that will be crowdfunded successfully.

5. Conclusions
This investigation contributes towards the understanding of the features of projects that influence crowdfunding success. No prior study has examined the effect of color combination within listing images on successful crowdfunding, nor the effect of socialness as a unified theoretical construct on successfully achieving funding.

Moreover, the research contributes a blueprint for a research process combining various elements and methods to explore image, text, and other features on behavioral outcomes. The research demonstrates the application of methods to measure theoretical concepts in text and images at a large scale and the use of these features in statistical analysis such as logistic regression to measure the probability of outcomes.

The findings from this research support the hypothesized relationships between both color complexity and socialness on successful project crowdfunding. The research demonstrates that crowdfunding listings are social artefacts, to which viewers may exhibit a level of parasocial relationship. From a practical viewpoint, the findings indicate that to increase the odds of successful crowdfunding, project developers can emphasize social cues in text descriptions and listing images for their project listings. Ideally, listing images should capture multiple human faces to increase the chances of successful crowdfunding, but at least one face will increase the odds of funding. This implies that carefully curated, high quality images (image quality was also found to be significant in our model) should be prominent in listings. Social words used in carefully developed listings that are not too long (length of description had a significant negative relationship with funding success) in a listing will also generate a marginal increase in the odds of successful funding.

The research brings a greater awareness to researchers of the important role that colors in combination have on individual psychology of website users. Previous research has focused on the effects of individual colors, but color combinations must be considered as a driver of individual attitudes and behaviors in research. In considering color complexity, the study finds that listing images containing fewer colors (palette simplicity) and that have more similar colors (higher color coherence or lower overall color distances) are more likely to be crowdfunded successfully. This finding is in line with principles from design practice and prior research by Barnes (2022), who found that lower color complexity in Airbnb accommodation listing images had a significant influence on visitor patronage. This implies that listing images must be screened carefully in order to increase the likelihood of appealing to potential backers, utilizing carefully selected, similar colors from a small palette, avoiding images with excess luminance (which was found to have a significant negative relationship with funding success). Such a process is likely to be more efficient if conducted via the assistance of professional graphic designers and/or systematic testing on a sample of test subjects.

The research has notable limitations. Firstly, the study was based on a data set from Kickstarter. Even though Kickstarter is the biggest platform according to the number of projects, many other crowdfunding platforms are available, such as Crowdfunder, Fundable,
GoFundMe, Indiegogo, StartEngine, Patreon, Crowdcube, and Mightycause. Future studies should attempt to test the hypotheses from this study using data from other platforms. Secondly, the Kickstarter platform is Western-centric, concentrating on North America, countries in Australasia and Europe, plus Singapore, Mexico, Hong Kong, and Japan. Future research should seek to investigate whether cultural differences exist in for socialness and color complexity and its relationship with funding success in and between a more varied group of countries. Finally, future research may seek to incorporate other variables in the analysis. For instance, emotion is not included in this study, but has been measured recently in other studies such as Hou et al. (2023). Future research may seek to assess emotion in faces detected in listing images, sentiment in descriptive text, and other additional features to evaluate their impact on successful crowdfunding.

References


International Journal of Human-Computer Studies, Vol. 43 No. 2, pp. 223-239, 


Effects of socialness and color complexity


About the author
Stuart John Barnes is Professor at Newcastle University Business School, Newcastle, UK. His research spans across academic disciplines, particularly information systems and marketing. His work has been published more than 200 times, including in leading outlets such as the Information Systems Journal, Information and Management, Tourism Management, Social Science and Medicine and Industrial Management and Data Systems, among many others. Professor Barnes is Fellow of the Royal Society of the Arts and the Royal Statistical Society. Stuart is rated among the top Business and Management researchers on Research.com (#67 UK, #405 World). He reviews for over 10 international grant awarding bodies. Stuart John Barnes can be contacted at: stuart.barnes@newcastle.ac.uk

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com