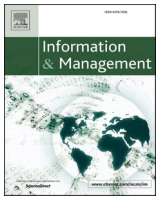
Information & Management 59 (2022) 103591

Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im



Linguistic understandability, signal observability, funding opportunities, and crowdfunding campaigns

Wei Wang [a], Yuting Xu [a], Yenchun Jim Wu [b,c,*], Mark Goh [d]

[a] College of Business Administration, Huaqiao University, Quanzhou City, Fujian
[b] College of Humanities and Arts, National Taipei University of Education, Taipei City, Taiwan
[c] Graduate Institute of Global Business and Strategy, National Taiwan Normal University, Taipei City, Taiwan
[d] NUS Business School and TLI-AP, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Crowdfunding
Text readability
Linguistic understandability
Information observability
Indiegogo

ABSTRACT

This study examines the impact of readability and understandability of the narratives on the signal receiver's acceptance of the signal which, in turn, affects a potential funder's decision, and hence the funding outcomes. Lexicon-related indicators and grammar-related indicators are applied to analyze 267,688 crowdfunding projects from the crowdfunding platform Indiegogo. Entrepreneurs must apply the right mix of lexicon- and grammar-related indicators to improve the online narrative's readability and understandability so as to send the right signals to attract project funders. Our results suggest that most of the indicators promote successful funding. Some indicators show a curvilinear, U-shape or inverted U-shape influence. The lexicon-related indicators rank higher in importance than the grammar-related indicators, and the number of difficult words ranks first among the indicators. This study links signal observability, linguistic cues, and funding success, and provides theoretical and practical implications on the understanding of signaling theory and narratives.

1. Introduction

Online entrepreneurs often apply the low-cost online financing model of crowdfunding to raise funds, donations, or solicit loans from potential backers in the crowd [1]. At the same time, the potential backers often use text as a proxy to assess the quality of a crowdfunding campaign [2]. Hence, the online entrepreneurs need to know how to effectively communicate their projects using appropriately pitched text descriptions to inform and educate the potential backers to persuade them to fund [3, 4]. This has to be balanced with the need to safeguard the project from being easily imitated or plagiarized [5]. Some theoretical lenses (such as signaling theory) highlight the importance of incomplete information from the marketplace that backers collect, organize, and evaluate for rational decision making, whereas other practices (such as patent protection) focus on enhancing understanding to deal with possible acts of plagiarism [6]. Most of these studies underline the antecedents of funding success, albeit ignore the balance between linguistic understandability and creative idea protection [7].

Signaling theory posits that certain factors determine an effective signal, and backers often rely on two aspects to judge signal quality: (1) the cost of the signal, and (2) whether the signal is effectively transmitted and communicated, and the receivers fully observe and understand it [8]. Researchers often care about whether the signal sender sends the signal, but pay little heed to whether the signal receivers understand the signal effectively. There is an opportunity to discuss the value of the signal only if the signal is observable. Hence, the signal observability of crowdfunding projects warrants research investigation in entrepreneurship studies.

As readability reflects the entrepreneur's educational level or social status, ceteris paribus, which informs that entrepreneurs who generate a higher readability text may be better educated [9]. Nevertheless, signal clarity does not always guarantee the funding success of a project. Put simply, text creators need to protect the project's intellectual property (IP) by deliberately reducing the understandability, which is clearly obvious in patent protection [10]. In an online crowdfunding model, entrepreneurs show their creative ideas through text description. This has led to numerous documented cases of plagiarism, copying of project ideas [11, 12]. Therefore, crowdfunding is a double-edged sword [13]. In many cases, entrepreneurs focus on the success of financing when describing a campaign, thus enhancing the observability and neglecting the protection of creativity, which makes it difficult to achieve the balance of "observability conversion" and "IP protection" [10], for this

* Corresponding author.
E-mail addresses: wayswang@gmail.com (W. Wang), 19013120022@stu.hqu.edu.cn (Y. Xu), wuyenchun@gmail.com (Y.J. Wu), bizgohkh@nus.edu.sg (M. Goh).

https://doi.org/10.1016/j.im.2022.103591
Received 12 November 2020; Received in revised form 1 January 2022; Accepted 4 January 2022
Available online 10 January 2022

reason, it has led to many IP disputes in crowdfunding [14]. There is thus a need to strike a balance between signal observability and safeguarding the project idea when generating the narratives [15].

Invariably, obfuscated and convoluted language is applied in corporate reporting either inadvertently or intentionally to mask the financial health [16]. Introducing unnecessary complexity into financial statements is a common tactic used to conceal and confuse [17]. This has led to annual reports becoming increasingly longer, more redundancy, ambiguous and convoluted, which is often used to cover up scandals such as Enron [18]. Although entrepreneurs hope to use the signals to obtain sufficient funds, they do not want their readers to "steal" their ideas. Thus, they may try to hide certain elements or reduce linguistic understandability and signal observability. However, it has not been established whether the online entrepreneurs have adopted a similar strategy in their crowdfunding pitches.

Prior studies inform that a crowdfunding campaign's presentation influences the pledge outcomes [19]. Thus, entrepreneurs should improve their project presentation, notably on the text description. However, the various aspects of readability have not been discussed and compared. That is, we lack an understanding and comparison of the two aspects of linguistic understandability, that is, lexicon-related and grammar-related indicators. The former focuses on lexicon analysis, and the latter on grammar mining [20]. We do not know how these two aspects promote linguistic understandability and safeguarding the project idea. Thus, we cannot offer enough guidance to the text creators.

The rest of the paper is set as follows. Section 2 reviews the theoretical foundation, summarizes the research gaps, proposes the research questions, and puts forward the hypotheses. Section 3 presents the research data, readability, and understandability detection model from the lexical and grammatical aspects. The findings are found in Section 4. Section 5 concludes with some discussion and future research directions.

2. Literature review and hypotheses development

2.1. Literature review

The extant studies have discussed crowdfunding extensively from various perspectives. Among them, the studies on the text description mainly start from the perspective of linguistic features, such as the influence of the subjectivity versus objectivity [21], text content emphasis [22], sentiment analysis [23, 24], fraudulent cues [25], linguistic styles [26, 27], psychological cues [28], and specific word list [29]. All of these studies show the value of narratives for successful financing. However, few studies focus on the readability of the description.

The narratives of crowdfunding projects on an open platform offer a variety of roles, that is, while attracting investors, it should also protect the idea [30]. Barrier-free text description is a cause for the endless creative idea plagiarism [31], and the incomprehensibility of patent documents is one of the means to safeguard creativity [32]. Signaling theory provides a basis for analyzing the linguistic features. In addition to the cost of the signal, efficient signal transmission requires that the signal be accurately sent, observed, and understood. Thus, signal quality depends heavily on the observability of the signal [33]. Research have noted the value of observability and understandability for signal transmission. Highly complex environments, such as the Internet, enhance the introduction of new ideas [34]. However, whether this idea can be safely "sold" to potential backers and be protected from plagiarism is an issue confronting the online entrepreneurs.

Unlike the offline financing channels, crowdfunding provides the entrepreneurs with the chance to attract their backers at a lower cost. Crowdfunding widens the channel of the entrepreneurs to reach their potential backers; this lifts the likelihood of funding success while reducing the financing cost [35]. The signal cost often determines the quality of the project, but on the Internet, the potential backers cannot rely on expensive signals alone, because the obstacles to using such financing tools are much lower. Many crowdfunding platforms provide sufficient applications (apps) for the entrepreneurs and backers. Another rarely discussed cost of signal transmission is the cost of preventing plagiarism. The cost of creative protection will increase if the contents of projects are displayed in detail [36]. Therefore, the impact of the entrepreneur's crowdfunding on venture capital investors is not always positive [37].

In cases when the signal cost does not render sufficient differentiation, signal observability becomes a problem for the backers. Although observability affords more opportunities, it also brings about greater challenges. Crowdfunding allows entrepreneurs to focus on clearly communicating their signals through the narratives [22]. Further, both parties must address the problem caused by the information asymmetry; the receiver has only the signals that the entrepreneur specifically chooses, to judge the prospects of the project. Many signals are often communicated ineffectively or incompletely, making them difficult to understand. Moreover, entrepreneurs may deliberately (and for good reason) highlight the good news and hide the bad. Thus, in an open environment with low signal costs [38], a signal's observability is valuable in the presentation of a crowdfunding project.

To see how readability can influence investors, some studies examine the specific linguistic attributes regarding understandability [39]. The contents will affect the composition, form, structure, meaning, and readability. These indicators have attributes that reveal the quality of the campaigns, by the intent of the signal transmitted by the entrepreneurs [33].

The readability and understandability of a signal, affecting the observability of the signal, have several effects. For one, readability, often denoted by a value [40–42], estimates the extent of a narrative that can be understood by a potential backer. Studies have shown the benefits of using readability in pedagogical applications: teachers often use such indicators to find and classify teaching resources; doctors can build an electronic medical record system with the help of text understandability [43]. However, the lack of structure and the brevity of online texts often render the standard natural language analysis techniques insufficient [44], which can challenge the processing of online crowdfunding text. From a funding perspective, clearly, the linguistic understandability and signal observability greatly improve the chances of successful financing, and there is no suspense about having a positive impact [45]. However, from another perspective, linguistic understandability and signal observability enhance the difficulty of creativity protection. As such, it impacts crowdfunding projects negatively [22].

2.2. Research gaps and questions

Summing the current research, at a minimum, the following gaps remain: (1) Text description is a factor that influences the pledge results of a crowdfunding project [46]. However, there is a dearth of studies on the role of readability and understandability on the pledge results, in the context of signal transmission and investment decisions. (2) In terms of the elements that make up the text, we still do not know the difference between the influence of the lexicon and grammar. Thus, we cannot balance the importance between the two aspects. (3) Similar to patent documents using obscure language to protect ideas, in an open online environment, it is unclear if entrepreneurs have adopted similar strategies. To fill this gap, we pose the following research questions:

1 How can readability and understandability be measured in crowdfunding campaigns? What is the form of the relationship between readability and the pledge results?
2 Readability encompasses both lexical and grammatical aspects. What is the difference in impact between these two types of indicators? Of all the readability indicators, which is more important?

2.3. Hypotheses development

The readability of the narratives conveys rich information of the

entrepreneurs and project. To the potential investors, the higher the readability and understandability of the project, the more observable is the project signal [20]. Investors are more likely to agree on the prospects of the campaign. Readability refers to the extent that the text can be understood by readers [41]. It may influence consumption intention positively [47]. Ceteris paribus, the entrepreneurs who generate higher readability texts are better educated [48], making them more trustworthy [49].

Generally, the description about a professional field tends to use some technical terms, which often reduces readability [32]. This is one reason why patent documents can sometimes be difficult to understand [50]. Therefore, entrepreneurs need to make a careful trade-off between readability and technical content. Too high a readability leads to lower technical content, and vice versa.

The literature reports that the influence of linguistic features presents an inverted U-shape [51, 52]. For example, an analysis of the online community found an inverted U-shaped relationship between users' past successful experience and idea implementation. Likewise, supporting evidence and idea implementation also show an inverted U-shaped relationship [53], as well as some linguistic indicators [26]. Therefore, we posit the following hypothesis:

Hypothesis 1. Linguistic understandability bears a curvilinear, inverted U-shaped influence on financing outcome.

Entrepreneurs have many choices to protect their project creativity, one of which is to reduce the readability, akin to patent protection [54]. The readability of a text can be categorized as lexicon and grammar. The former concerns words; the latter focuses on the grammar [55].

Readability is an important linguistic feature which determines whether the audience can accurately identify the lexical meanings. However, once they encounter the difficult words they do not acknowledge, the reader cannot fathom the meaning, which is much more influential than the complex grammar [51]. In the linguistic research in crowdfunding, many studies analyze the subjectivity versus objectivity aspects [21], sentiment [24], semantics [23], and topics [56]. All of these studies are conducted from the perspective of the lexicon rather than the grammar. In short, these studies move that the influence of grammar on readability is less important than that of the lexicon when analyzing crowdfunding narratives. Hence, the lexicon-related indicators appear to be more important than grammar-related readability indicators in the influence of the readability and understandability. Thus, we posit the following hypothesis:

Hypothesis 2. Lexicon-related indicators rank higher in importance than grammar-related indicators in crowdfunding linguistic understandability.

3. Research data, text mining model, and empirical models

3.1. Research data

Many extant research focus on the online platform Kickstarter which adopts the *All-or-Nothing* funding mode, and the funding outcome is either a success or failure. If the pre-determined goal is unmet, the funding project will be deemed unsuccessful and the money will be returned to the funders. In some cases, the *All-or-Nothing* funding model shows weaknesses, such as poverty alleviation, education, and environmental protection [57]. Another crowdfunding platform, Indiegogo, adopts an *All-and-More* funding model, whereby even if the pre-determined goal is not met, the entrepreneur will receive the cumulative sum of money, at the cost of a higher platform fee. This funding mode is very popular for online projects, such as education support for the under-privileged, medical support for rural areas, and protection of living things [57]. Even if the capital raised cannot meet the target, a marginal increase in funds can allow more folks to gain better social welfare or provide slightly better protection of animals and plants. In view of this aspect of Indiegogo, we choose it as the research object.

We developed a crawler with Python to track the crowdfunding data, collecting 267,688 public campaigns from Indiegogo, as shown in Table 1. These campaigns exclude the non-public display projects and non-English projects. Among the categories, the success rate of "theater" is the highest (25.24%), whereas "small business" is the lowest (2.85%).

Fig. 1 shows the descriptive statistics for the sample. Most of the campaigns draw less than 20 backers; specifically, the mean number of backers in the sample is 14. In terms of the pledge money, though most of the projects pledged a relatively small amount of money, some projects have pledged a very large amount of money, and thus the average amount pledged is $1455. On financing progress, most projects are under 20%, with some between 100% and 120%; the backers' investment momentum will drop significantly when the financing progress reaches 100% [58]. Finally, the mean funding success rate of Indiegogo projects is 9.12%, compared to Kickstarter's success rate of 37%. This is due to the difference in the funding mechanisms of Indiegogo and Kickstarter.

3.2. Readability and understandability detection model

We adopt a data analytics approach, following the extant literature. Ten indicators are used to measure text readability and understandability [40–42]. Table 2 summaries the operational definitions of text readability and understandability.

The following measures are used as lexicon-related indicators:

(1) The Flesch Reading Ease [59] estimates a Flesch score as shown in Eqn (1), which counts both the *words per sentence* and *syllables per word*. The larger the number, the better is the readability, that is, 100 means very easy, whereas 0 indicates very confusing.

$$FRE = 206.83 - (1.015 * wordspersentense) - (84.6 * syllablesperword) \quad (1)$$

(2) The Coleman–Liau index [60] returns the grade level of the text using the Coleman–Liau formula, as shown in Eqn (2), where, in a sample string of 100 words, w is the number of one-syllable words found, s indicates the number of sentences, p represents the number of pronouns, and $prep$ is the number of prepositions found, respectively.

$$C\% = 104w + 1.06s + 0.56p - 0.36prep - 26.01 \quad (2)$$

(3) The Linsear Write Formula [61] returns a grade level, which is found by the length of a sentence as well as the words containing 3 or more syllables as shown in Eqn (3), where r is the point score for a 100-word sample (1 point for an easy word and 3 points for a hard word). For example, a score of 9.2 means that a ninth grader would be able to read and comprehend the document.

$$\begin{cases} LW = r/2, \text{If } r > 20 \\ LW = r/2 - 1, \text{If } r \leq 20 \end{cases} \quad (3)$$

(4) The difficult words indicator [62] assumes that the difficulty in text reading is due to the proportion of difficult words in a narrative; difficult words are those words other than the 3000 commonly used words.

The following measures are used as grammar-related indicators:

(1) The Automated Readability Index (ARI) estimates the grade level that a reader must have to understand the text [60] as shown in

Table 1
Sample statistics of categories.

Category	Quantity	Percentage	Goal ($)	Pledge ($)	Progress (%)	Backers count	Success rate (%)
Animals	8367	3.13	77,764.61	732.76	17.38	10.60	7.60
Art	8380	3.13	85,371.83	1179.80	27.20	16.94	12.37
Comic	1763	0.66	346,652.30	1405.21	29.21	19.52	13.56
Community	38,706	14.46	159,927.30	976.69	20.50	11.42	8.30
Dance	2462	0.92	9334.48	1406.49	37.43	18.75	18.97
Design	3585	1.34	39,869.98	3055.02	23.09	16.90	6.75
Education	22,869	8.54	79,791.98	901.78	16.65	10.34	7.88
Environment	4663	1.74	394,954.30	1225.82	14.16	13.63	6.71
Fashion	4669	1.74	29,045.35	1450.72	23.83	12.48	6.06
Film	40,229	15.03	91,163.15	1848.47	24.65	18.97	11.87
Food	5296	1.98	110,752.70	1295.61	12.40	15.33	5.65
Gaming	7638	2.85	110,516.50	989.26	15.90	11.96	5.11
Health	20,085	7.50	89,357.33	1149.75	19.80	12.96	7.94
Music	20,120	7.52	35,718.54	1528.99	29.92	23.59	15.17
Photography	4389	1.64	11,603.77	807.18	16.93	10.61	7.75
Politics	2299	0.86	667,103.10	1680.55	19.71	19.69	9.66
Religion	3712	1.39	189,172.60	714.99	11.53	6.13	4.96
Small Business	21,932	8.19	173,165.20	602.72	6.55	7.00	2.85
Sports	6231	2.33	674,180.70	790.07	17.82	9.61	8.14
Technology	14,325	5.35	352,306.10	5368.19	26.86	14.54	4.65
Theater	7723	2.89	9414.74	1930.97	47.29	26.58	25.24
Transmedia	1251	0.47	158,620.60	1728.02	19.49	19.59	8.79
Video / Web	9324	3.48	81,222.69	1555.79	16.86	14.27	7.69
Writing	7638	2.85	15,699.15	846.42	20.55	14.55	9.79
Others	32	0.01	8005.63	320.00	10.59	3.31	6.25
Sum	267,688	100	135,335.60	1455.50	21.05	14.39	9.12

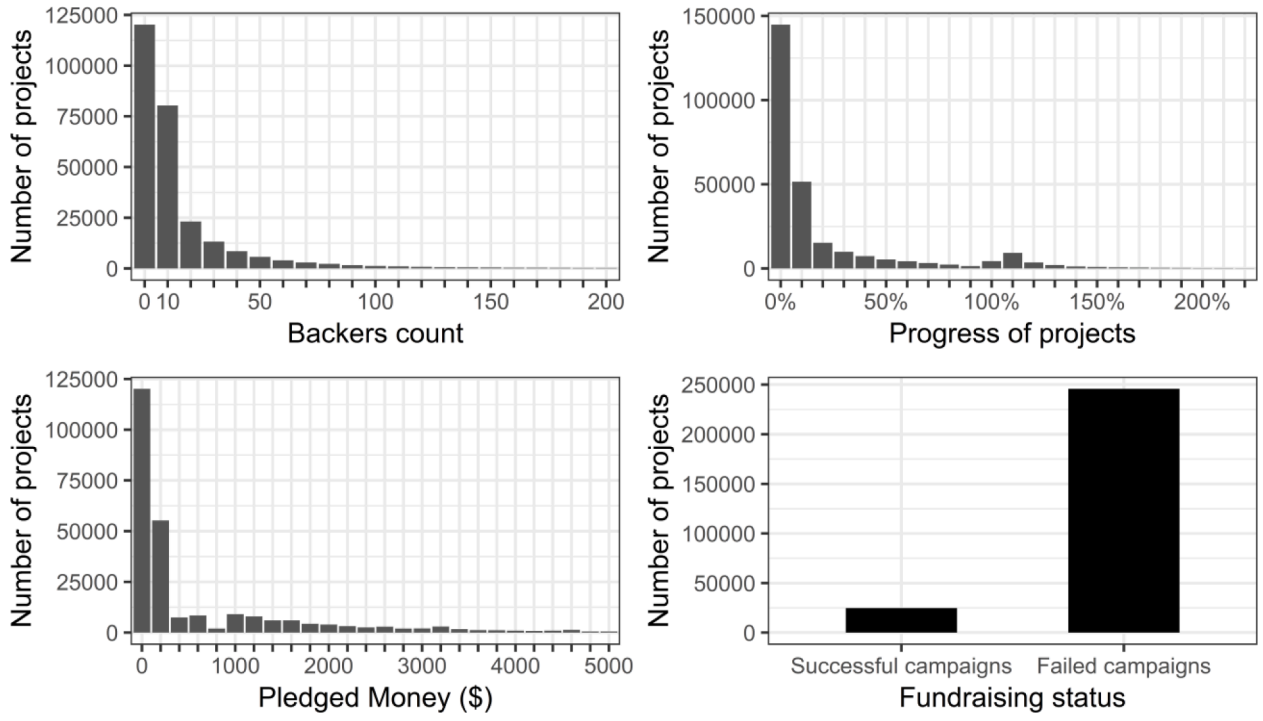


Fig. 1. Statistical profile of the sample.

Eqn (4), where c is the number of letters and numbers, w is the number of spaces, and s is the number of sentences, respectively. For example, if the ARI is 6.5, then the reader should be a grade 6 to grade 7.

$$ARI = 4.71 \times \left(\frac{c}{w}\right) + 0.5 \times \left(\frac{w}{s}\right) - 21.43 \quad (4)$$

(2) The SMOG index [63] returns the SMOG score of a given text as shown in Eqn (5), which counts the number of words of 3 or more syllables (*polysyllable count*) in the proper number of sentences. As long words and sentences can cause poor readability, the SMOG index considers both word length and sentence length.

$$SMOG = 3 + \sqrt{polysyllable\ count} \quad (5)$$

(3) The Flesch-Kincaid readability score [64] returns a grade score obtained by the Flesch-Kincaid grade formula as shown in Eqn (6), where *ASL* is the average length of the sentences, and *ACW*

Table 2
Definitions of textual readability and understandability.

Classification	Indicator	Definition
Lexicon-related indicators	Flesch Reading Ease [59]	Scores material by considering (1) the total number of words, (2) the number syllables, and (3) the total number of sentences.
	Coleman–Liau index [60]	Scores material by words, syllables, pronouns, and prepositions.
	Linsear Write Formula [61]	Scores material based on monosyllabic words and strong verbs which use in official writing aimed at the public and U.S. navy.
	Difficult words [62]	Measures whether the word is a common word.
Grammar-related indicators	Automated Readability Index [60]	Measures the characters used to assess the U.S. grade level required to read a piece of text.
	SMOG index [63]	Measures how many years of education an average person needs to have to understand a text.
	Flesch-Kincaid Readability score [64]	Use a table to translate to the reading grade level which assesses the approximate reading grade level of a text.
Mixed indicators	Gunning Fog Index [65]	Estimates the education level required to understand the text.
	Dale-Chall readability score [66]	Measures a text against the words and sentences considered familiar to fourth-graders.
	Text standard readability consensus [67]	A comprehensive indicator considering the vocabulary, grammar, words and sentences, which returns the estimated school grade level required to understand the text.

denotes the average number of syllables. For instance, a score of 8.0 suggests that an eighth-grade student can read and understand the text.

$$FKR = 0.39 \times ASL + 11.8 \times ASW - 15.59 \tag{6}$$

The following measures are used as mixed indicators:

(1) The Gunning Fog Index [65] returns the Fog value of the given text estimated by the Gunning Fog formula as shown in Eqn (7), where *ASL* indicates average length of the sentences, *ACW* denotes the average number of syllables, *ACW* is the average number of complex words per 100 words (words that contain two syllables or more). The smaller the value, the less is the difficulty of the words.

$$FOG = 0.4 \times (ASL + 100 \times ACW) \tag{7}$$

(2) The Dale-Chall readability score [66] uses a list of 3000 commonly used English words and returns a grade level using the Dale-Chall formula as shown in Eqn (8), with *dw* as the number of difficult words (out of a list of 3000 commonly used words), *w* as the number of words, and *s* as the number of sentences, respectively. This formula is specially designed as it adopts the "hard" words accepted by grade 4 students.

$$DCRS = 0.1579 \times \left(\frac{dw}{w} \times 100\right) + 0.0496 \times \left(\frac{w}{s}\right) \tag{8}$$

(3) The text standard readability consensus [67] returns the best grade level of a given text, as shown in Eqn (9).

$$TS = \frac{(FRE + CLI + LW + DW + ARI + SMOG + FKR + FOG + DCRS)}{9} \tag{9}$$

In terms of technology implementation, Python 3.7.4 with *NLTK* 3.4.5 and *textstat* 0.5.4 are used to pre-process the text and estimate the ten indicators of text readability and understandability, and all output are in numerical form.

3.3. Empirical models

Indiegogo adopts the *All-and-More* funding model, which possesses certain unique features. The entrepreneurs positioning their projects on Indiegogo typically do so for reasons which include: (1) capital requirement [68]; (2) getting more participants, for example, some projects need sufficient participants to label data or collect local biological data [69, 70]. The entrepreneurs set a funding target when launching the campaign. When the target is met, the platform's service charge is significantly reduced; otherwise a higher service charge is meted [57]. On Indiegogo, even if the funding target is achieved, investors can continue to invest in the projects till the end of the investment duration, which leads to the financing progress of many projects being more than 100% (i.e., the actual amount funded is greater than the amount planned) [71].

In view of the above analysis, to measure the influence of linguistic understandability from multiple perspectives, four measurements are used as dependent variables: (1) dummy variable, the dependent variable is set to 1 if the pledged money exceeds the target, otherwise it is 0 [22]; (2) the ratio of pledged money [21]; (3) the number of backers [72]; and (4) the pledged money [73]. Next, four empirical models, drawn from the works mentioned, are built to estimate the effect of readability and understandability among the scenarios with four dependent variables of *Status* (1=*success*), *NumProgress* (%), *NumBackers*, and *NumPledgedMoney*, respectively.

Model 1: Status model (dummy variable) is shown in Eqn (10), where $Status_i$ is 1 if the campaign is funded successfully, and 0 otherwise. $R_i^{'}$ represents the readability and understandability of project *i*, and $C_i^{'}$ are the control variables, respectively. ε_i is the random disturbance factor, which is assumed to follow $\varepsilon = N(0, \delta^2)$. α is the intercept term, whereas β and γ are the coefficients for the readability and the control variables, respectively.

$$Status_i = \alpha + R_i^{'} \times \beta + C_i^{'} \times \gamma + \varepsilon_i \tag{10}$$

Model 2: The progress model follows Eqn (11), where $Progress_i$ is the pledged ratio of project *i* (funds raised divided by financing targets).

$$\log(Progress_i) = \alpha + R_i^{'} \times \beta + C_i^{'} \times \gamma + \varepsilon_i \tag{11}$$

Model 3: The backers model follows Eqn (12), where $Backers_i$ is the number of backers for project *i*.

$$\log(Backers_i) = \alpha + R_i^{'} \times \beta + C_i^{'} \times \gamma + \varepsilon_i \tag{12}$$

Model 4: The pledge money model follows Eqn (13), where $Money_i$ is the amount pledged for project *i*.

$$\log(Money_i) = \alpha + R_i^{'} \times \beta + C_i^{'} \times \gamma + \varepsilon_i \tag{13}$$

Next, with Stata 16, we apply each of the four models to assess the effect of readability and understandability on the status of the projects, the progress of project funding, the number of backers for the projects, and the pledge dollar amount for the projects reported on Indiegogo.

4. Results

4.1. Statistical results

Table 3 presents the sample statistics. From prior studies [21, 26, 27, 74], the models have 6 control variables (see project related category in Table 3). *NumTeamFounder* is the size of the entrepreneurial team initiating the crowdfunding project [75]. A mean value of 1.5 informs that the most campaigns on Indiegogo are launched and maintained by at least one entrepreneur. *NumPledgeLevel* indicates the number of levels of investment amounts that the backers can choose from. Naturally, different investment amounts yield different rewards. On the whole, entrepreneurs on Indiegogo set an average of four participation levels (4.323 in value).

For the readability and understandability indicators, the lowest average score of the 10 readability indicators is the Dale-Chall readability score (*NumDCRS*) with 7.545, whereas the highest is the difficult words indicator (*NumDW*) with 77.317.

4.2. Effects of readability and understandability on financing outcomes

Model 1 is fitted by a logit model, and Models 2 to 4 are estimated using least squares regression. Table 4 shows the effects of readability and understandability on the financing outcomes, which shows significant differences among the indicators. Some variables show different, or even diametric, effects among the models. First, *NumFRE* estimates the Flesch score, where the larger the score, the better is the readability. In all of the models, the regression coefficient is positive impacted and statistically significant, implying that the readability of the Flesch score promotes successful funding. Similar indicators include *NumCLI, NumDCRS*, and *NumGF*. As *NumDCRS* indicates difficulty, a larger score indicates poorer readability and understandability; thus, the negative regression coefficients indicate that readability and understandability also promote good funding outcomes.

For *NumLWF*, the regression coefficient in the Model 4 is 0.0803***, as it considers both the length of the sentences and words [61], with the greater the value, the poorer is the readability. Thus, the positive influence indicates that some readability and understandability indicators will lessen the likelihood of successful financing. In short, text readability and understandability are not always positive on financing. This finding highlights the challenge faced by investors when evaluating the narratives. Therefore, entrepreneurs must apply a basket of readability indicators when generating the narratives to ensure that the project signal is accurately transmitted to the potential backers.

4.3. Marginal effects of readability and understandability

With the Status model, logit regression is applied to estimate the marginal effects of readability and understandability. Regression analysis helps one to understand how the dependent variable changes with an independent variable, whereas the marginal effects measure the percentage change in the financing outcome when the readability and understandability shift by 1% [76]. Table 5 shows the marginal effects.

The marginal effects are not always consistent with the main effect. The marginal effects of readability and understandability show some differences among the indicators. The most influential factor is *NumCLI* (marginal effect = 0.0413***), and the least affected is *NumTS* (marginal effect = –.0006). Thus, entrepreneurs should heed the Coleman–Liau result; they should use simple words and short sentences to describe their projects, to ensure a higher likelihood of funding success.

4.4. Readability and understandability's inverted U-shape effect

The impact of readability and understandability is non-linear, bearing a U-shape or inverted U-shape relationship. This would suggest that, on some indicators, there is a threshold, and there are clear differences on the influence before and after the threshold [21, 77]. As Indiegogo uses the *All-and-More* funding model, we use the pledge money model to illustrate the coefficient of the quadratic terms. The results are shown in Fig. 2. We note that *FRE* means that the larger the value, the higher is the readability of the text; also, the larger the value of *ARI*, the lower is the readability. Therefore, the U-shape impact of *FRE* is equivalent to the inverted U-shape influence of *ARI*. Hence, Hypothesis 1 is supported.

When the text creator of a project narrative introduces professional content to the narrative, it is inevitable to use some technical verbiage. The readability of the technical terms is usually low such as those found in patent and legal documents [78]. However, this rule is not completely true in crowdfunding. Some readability indicators have a long rising curve, such as *LWF*; however, other indicators have very short rising curve, such as *ARI* and *DCRS*.

4.5. Importance ranking for signal observability

Compared to the explanatory models, predictive models are more practical as the entrepreneurs are more interested in knowing the possibility of successful online financing when launching a crowdfunding campaign, and to mitigate the possibility of plagiarism. They also want to know what aspects should be improved if they were to modify the text

Table 3
Descriptive statistics (N = 267,688).

Category	Variable	Description	AVG	S.D.	Median	Min.	Max.
Pledge results-related	*Status (1=success)*	Pledge results	.091	.288	0	0	1
	NumBackers	Count of backers	14.394	43.835	1	0	996
	NumProgress (%)	Pledge progress	21.053	119.500	0	0	20,705
	NumPledgedMoney	Pledge money	1455.497	29,618.67	20	0	1.28e+07
Project related	*NumUpdates*	Update count	1.581	5.040	0	0	498
	NumComment	Comment count	5.880	21.398	2	0	982
	NumGoal	Pledge goal	135,335.6	2,764,305	5000	0	1e+08
	NumLastdays	Pledge duration days	48.615	27.847	45	0	365
	NumTeamFounder	Count of team member	1.525	1.456	1	1	125
	NumPledgeLevel	Pledge levels count	4.323	3.629	4	0	56
Readability related	*NumFRE*	Flesch reading ease	62.475	22.870	65.35	−3842.13	148.06
	NumSMOG	Smog index	10.114	3.728	10.7	0	40.2
	NumFKG	Flesch Kincaid grade	9.316	4.601	8.7	−5.5	550.4
	NumCLI	Coleman–Liau index	11.669	69.394	11.15	−39.61	30,926.9
	NumARI	Automated Readability Index	11.942	56.536	11.1	−16.3	25,119.2
	NumDCRS	Dale Chall score	7.545	1.238	7.47	.05	27.83
	NumDW	Difficult words	77.317	83.430	53	0	1640
	NumLWF	Linsear write formula	8.451	4.761	7.333	−1	64
	NumGF	Gunning Fog	17.241	4.500	16.776	2.4	187.012
	NumTS	Text standard	9.560	4.836	9	−3	447

Table 4
Effects of readability and understandability on pledge results.

Variable	Status model		Backers model		Progress model		Money model	
	1	2	1	2	1	2	1	2
NumUpdates	.213***	.1894***	.1512***	.1431***	.1234***	.1176***	.1643***	.1562***
	(0.01)	(0.0102)	(0.0037)	(0.0037)	(0.0044)	(0.0044)	(0.0055)	(0.0055)
NumComment	1.15***	1.158***	.7412***	.7424***	.7235***	.7285***	.9888***	.9894***
	(0.0104)	(0.0106)	(0.0033)	(0.0033)	(0.004)	(0.004)	(0.0049)	(0.0049)
NumGoal	−1.039***	−1.105***	−0.0947***	−0.1077***	−0.6739***	−0.7008***	−0.0473***	−0.0656***
	(0.0096)	(0.0102)	(0.0023)	(0.0024)	(0.0034)	(0.0035)	(0.0034)	(0.0035)
NumLastdays	−0.4298***	−0.4328***	−0.2112***	−0.2121***	−0.2903***	−0.2906***	−0.2808***	−0.2881***
	(0.0152)	(0.0157)	(0.006)	(0.006)	(0.0072)	(0.0072)	(0.009)	(0.0089)
NumTeamFounder	.3787***	.3114***	.2653***	.2305***	.2343***	.1985***	.4144***	.3478***
	(0.0148)	(0.0152)	(0.0054)	(0.0054)	(0.0064)	(0.0065)	(0.008)	(0.0081)
NumPledgeLevel	.5023***	.3888***	.3318***	.2995***	.2444***	.2102***	.4061***	.35***
	(0.0176)	(0.0187)	(0.0057)	(0.0061)	(0.0072)	(0.0076)	(0.0086)	(0.009)
NumFRE		.4614***		.1612***		.1708***		.3584***
		(0.083)		(0.025)		(0.0318)		(0.0371)
NumSMOG		.3087**		−0.0403		.1588***		.3376***
		(0.1102)		(0.0374)		(0.0463)		(0.0554)
NumFKG		−0.19		−0.209***		−0.1564**		−0.2053***
		(0.1177)		(0.0395)		(0.0496)		(0.0586)
NumCLI		1.013***		.6437***		.821***		1.157***
		(0.1232)		(0.0426)		(0.0549)		(0.0631)
NumARI		−0.1755		−0.3024***		−0.4289***		−0.7239***
		(0.1494)		(0.051)		(0.0657)		(0.0755)
NumDCRS		−0.4973*		−0.3482***		−0.4974***		−0.4031**
		(0.233)		(0.0887)		(0.1182)		(0.1315)
NumDW		.2244***		.0822***		.0792***		.1039***
		(0.0122)		(0.0041)		(0.0051)		(0.0061)
NumLWF		−0.0539		.0176		.0214		.0803***
		(0.0305)		(0.0105)		(0.0128)		(0.0156)
NumGF		.9709***		1.015***		.9963***		1.491***
		(0.2474)		(0.09)		(0.1179)		(0.1334)
NumTS		.0156		.0014		.0245		.0133
		(0.0386)		(0.0135)		(0.0164)		(0.0199)
Constant	4.987***	−1.119	1.935***	−1.385***	7.663***	3.998***	4.968***	−1.667***
Adjusted R^2	0.322	0.329	0.467	0.478	0.440	0.453	0.407	0.425

* $p<0.05$, ** $p<0.01$, *** $p<0.001$, Standard errors in parentheses.

Table 5
Marginal effects of readability for the Status Model.

Variables	Main	dy/dx	Main	dy/dx
NumUpdates	.213***	.0086***	.1894***	.0077***
	(0.01)	(0.0004)	(0.0102)	(0.0004)
NumComment	1.15***	.0462***	1.158***	.0472***
	(0.0104)	(0.0006)	(0.0106)	(0.0006)
NumGoal	−1.039***	−0.0418***	−1.105***	−0.045***
	(0.0096)	(0.0004)	(0.0102)	(0.0005)
NumLastdays	−0.4298***	−0.0173***	−0.4328***	−0.0176***
	(0.0152)	(0.0006)	(0.0157)	(0.0007)
NumTeamFounder	.3787***	.0152***	.3114***	.0127***
	(0.0148)	(0.0006)	(0.0152)	(0.0006)
NumPledgeLevel	.5023***	.0202***	.3888***	.0158***
	(0.0176)	(0.0007)	(0.0187)	(0.0007)
NumFRE			.4614***	.0188***
			(0.083)	(0.0034)
NumSMOG			.3087**	.0126**
			(0.1102)	(0.0045)
NumFKG			−0.19	−0.0077
			(0.1177)	(0.0048)
NumCLI			1.013***	.0413***
			(0.1232)	(0.005)
NumARI			−0.1755	−0.0072
			(0.1494)	(0.0061)
NumDCRS			−0.4973*	−0.0203*
			(0.233)	(0.0095)
NumDW			.2244***	.0091***
			(0.0122)	(0.0005)
NumLWF			−0.0539	−0.0022
			(0.0305)	(0.0012)
NumGF			.9709***	.0396***
			(0.2474)	(0.0101)
NumTS			.0156	.0006
			(0.0386)	(0.0016)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$, Standard errors in parentheses.

description. Improving signal observability from the right aspect is thus imperative.

We therefore apply machine learning to evaluate the importance of the ten readability and understandability indicators adopted in this study. Cross-validation is a data partitioning method commonly used in machine learning. Following the literature, we use a 10-fold cross-validation on the predictive model. The random forest method is applied on Model 1 [79], and the linear regression algorithm is adopted for Models 2 to 4. The Python package *sklearn* 0.21.3 is applied to implement the machine learning algorithm. Table 6 shows the importance ranking for signal observability. All the importance scores are scaled on [0,100]. Note that an importance of 0 does not mean that a variable has no impact on funding outcome. Rather, that variable exerts less influence than the other variables.

In terms of importance, the difficult words indicator (*NumDW*) ranks first in all the models, that is, difficult words are the most important factor to predict the pledge outcomes. Therefore, entrepreneurs should heed that using the right choice of words in the text description (lexicon-related indicator) is crucial for communicating to the potential backers and win them over. We note too that the variables with the highest importance rank are almost related to the lexicons, whereas the variables related to the grammar rank less important. Thus, Hypothesis 2 is supported.

5. Discussion and contributions

5.1. General discussion

The potential backers, being rational, will always gather information and weigh the consequences of the alternatives, make decisions and take prudent action, which stresses the importance of sending clear signals

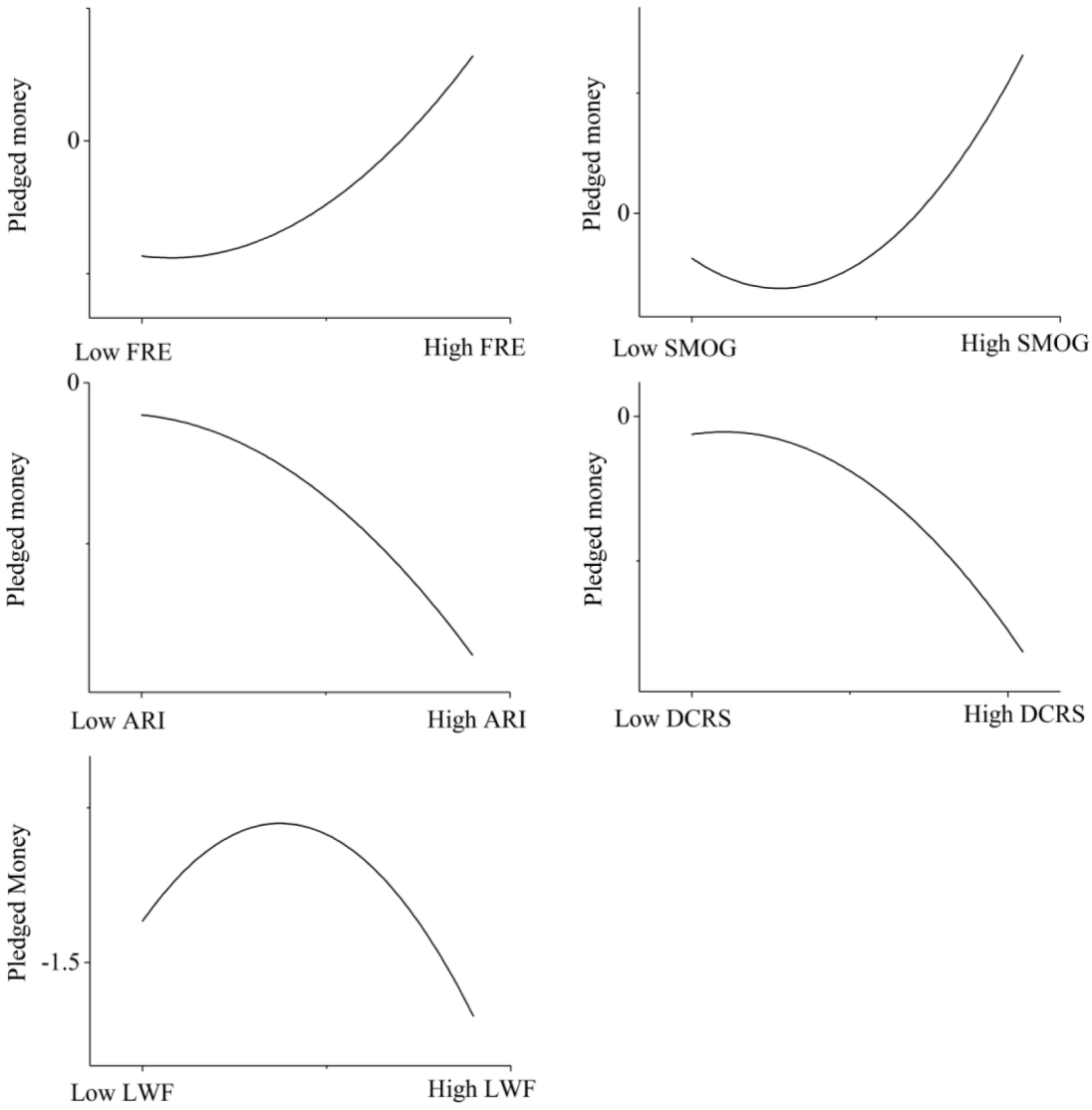


Fig. 2. Readability and understandability's inverted U-shape effect.

Table 6
Importance ranking for signal observability.

Variable	Importance Score (rank)			
	Status model	Backers model	Progress Model	Money model
NumDW	100 (1)	100 (1)	100 (1)	100 (1)
NumCLI	67.603 (2)	0.04402 (9)	0.05821 (9)	0 (10)
NumFRE	36.544 (3)	8.34598 (4)	19.28939 (3)	8.3902 (5)
NumLWF	13.518 (4)	16.32636 (2)	26.42838 (2)	9.5931 (3)
NumARI	9.154 (5)	0 (10)	0 (10)	0.01388 (9)
NumGF	6.638 (6)	0.52166 (7)	2.96703 (7)	6.14702 (6)
NumDCRS	6.565 (7)	1.29009 (6)	3.06052 (6)	1.78145 (8)
NumSMOG	6.019 (8)	0.22986 (8)	8.41675 (5)	12.0448 (2)
NumFKG	3.551 (9)	4.5358 (5)	2.80424 (8)	11.58844 (4)
NumTS	0 (10)	10.1926 (3)	14.77971 (4)	5.42131 (7)

Values in parentheses indicate the order of importance of the variable.

for informed decision making. However, there is a lack of systematic research on signal transmission in the realm of online financing. Sometimes, readability and technical content are a pair of contradictions. In the trade-off between readability and technical content, many studies have found that the influence of the linguistic features yields an inverted U-shape effect [26]. Our study reports that readability has a similar inverted U-shape effect, and there is a threshold to reverse the influence. In short, entrepreneurs cannot use too simple a textual description. They must balance text readability and technical content.

Furthermore, we divide the readability of text into two categories: lexicon-related indicators and grammar-related indicators. The former measures the difficulty of the words, whereas the latter measures the grammatical structure of the description. In the extant studies, few studies have considered the difference between these two types of readability indicators. Through importance ranking analysis and cross validation, this study finds that the lexicon-related indicators are more important than the grammar-related indicators. Therefore, entrepreneurs should give priority first to the words used, followed by grammar.

5.2. Contribution to signaling theory

The essence of communication is to make the signal receiver understand without bias the intended meaning of the signal from the transmitter, engendering a prerequisite for effective signal transmission [80]. If the signal transmitted is inaccurate, obscure, ambiguous or cannot be understood by the signal receiver unbiased, then the communication is deemed to be erratic. The signal transmission problem is particularly acute in an online setting for several reasons: (1) Information overload on the Internet. The potential backers may be presented with many similar projects and the current market characteristics

suggest that backers suffer from information overload and spending time to learn about the campaigns is costly. Therefore, once the potential backers find it difficult to understand, they are more likely to lose interest. (2) Backers decide on whether the entrepreneur is credible by observing the signal [81]. If an entrepreneur's narrative contains veiled explanations, incomprehensible content or misleading narratives, the signal receiver is likely to perceive the project as fraudulent [82].

Creative ideas cannot guarantee funding success if the transmitted signal is inaccurate and cannot be properly discerned by others, and most entrepreneurs are always challenged by insufficient funds [46, 83]. Therefore, raising as much funds as possible is crucial for the *All-and--More* funding model. Moreover, in order to protect the creativity, entrepreneurs may deliberately increase the difficulty of understanding the text, at the risk of poor project funding. With the help of signal interpretation on signaling theory [33], the contribution of this paper to signaling theory is now summarized in Fig. 3.

Thus far, much attention has been paid to the cost of the signal, than on the observability of the signal. In the early stages of entrepreneurship, entrepreneurs can only describe the outline of an idea, and it is difficult to articulate the details. Therefore, a scripted project narrative as a form of signal clarity should clearly and explicitly link the current situation and the vision for the future [84]. Studies have used various signaling approaches to woo backers. This study has focused on the readability and understandability of the textual description to improve signal clarity. With the ubiquity of crowdfunding and the cost of raising funds, our insights into signal observability can help entrepreneurs to understand how potential backers assess project quality.

Our thesis is that using appropriate signals to improve the readability and understandability in crowdfunding campaign descriptions will attract backers. However, this presents a challenge to safeguarding the entrepreneur's creativity. The readability and understandability of the narratives and the statements they embody reveal the nuanced quality of the entrepreneur's creative ideas. The scripted narratives are a subtle way of revealing advantages over the other campaigns. However, not all statements are conducive to readability and understandability. Investors tend to reject these ambiguous signals, which is why most of the indicators that are not conducive to signal observability have an adverse effect on project funding outcomes.

More importantly, to ensure the creativity of an online project, using a less straightforward language expression is a method that many text creators adopt. Many theories and practice assert that an effective mode of signal transmission is to consciously fuzz the key signals to attenuate the readability and understandability [7, 32]. In signaling theory, this study can thus serve to improve the aspects of creativity protection in signaling, which is a reference for applying signaling theory to IP protection and online financing.

5.3. Implications for entrepreneurial narratives

The scripted narratives are a subtle way of revealing the advantages over the other campaigns. However, not all statements are conducive to readability and understandability. Investors tend to reject those ambiguous signals, which is why most of the indicators that are not conducive to signal observability have an adverse effect on the project funding outcomes. The good news for entrepreneurs is the inverted U-shape effect of readability, which can delimit the boundaries of successful funding and safeguarding entrepreneurial creativity. In short, entrepreneurs need to cut a delicate trade-off between successful funding outcomes and creative idea leakage. The zenith of the inverted U-shape curve provides guidance for entrepreneurs to describe the projects. Further, our investigation suggests that the influence of lexical understandability outweighs that of grammatical understandability, thus lending guidance for entrepreneurs when generating their project narratives.

5.4. Limitations and further research opportunities

This study offers a framework and empirical models for signal observability in a crowdfunding environment, a relevant yet understudied topic. Future directions for research include: (1) platforms that have adopted other models, such as the *All-or-Nothing* funding model used by Kickstarter, (2) other factors affecting the effectiveness of signal transmission, such as the promised reward quantum, and (3) the project and funding mechanisms.

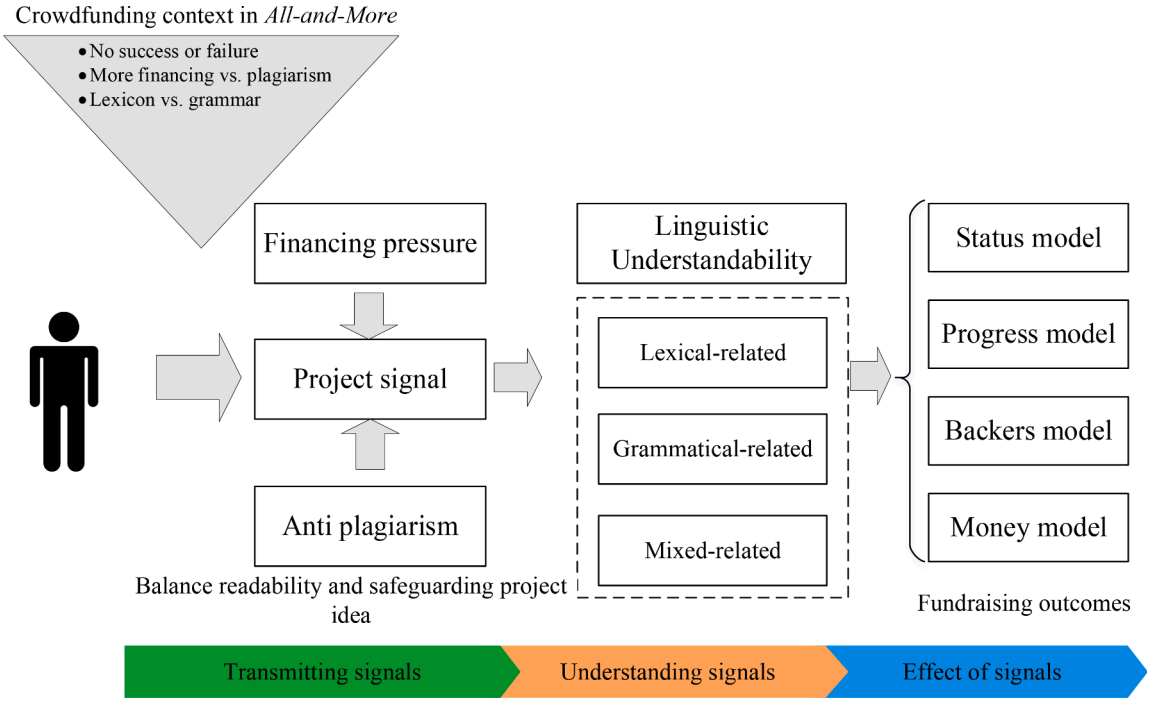


Fig. 3. Contribution to the signaling theory.

Declaration of Competing Interest

The authors have no conflicts of interest to declare

Acknowledgments

This work is partially supported by the National Nature Science Foundation of China Grant (72072062), and Natural Science Foundation of Fujian Province (2020J01782).

References

[1] S. Testa, K.R. Nielsen, M. Bogers, S. Cincotti, The role of crowdfunding in moving towards a sustainable society, Technological Forecasting and Social Change 141 (2019) 66–73.
[2] L. Mastrangelo, S. Cruz-Ros, M.-.J. Miquel-Romero, Crowdfunding success: the role of co-creation, feedback, and corporate social responsibility, International Journal of Entrepreneurial Behavior & Research 26 (2020) 449–466.
[3] P.H. Kim, M. Buffart, G. Croidieu, TMI: signaling credible claims in crowdfunding campaign narratives, Group & Organization Management 41 (2016) 717–750.
[4] Y.-.M. Li, J.-.H. Liou, Y.-.W. Li, A social recommendation approach for reward-based crowdfunding campaigns, Information & Management 57 (2020), 103246.
[5] T. Wang, X. Liu, M. Kang, H. Zheng, Exploring the determinants of fundraisers' voluntary information disclosure on crowdfunding platforms: a risk-perception perspective, Online Information Review 42 (2018) 324–342.
[6] C. Justicia De La Torre, D. Sánchez, I. Blanco, M.J. Martín-Bautista, Text mining: techniques, applications, and challenges, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 26 (2018) 553–582.
[7] M. Bukowski, S. Geisler, T. Schmitz-Rode, R. Farkas, Feasibility of activity-based expert profiling using text mining of scientific publications and patents, Scientometrics (2020) 1–42.
[8] M. Spence, Job market signaling, Q. J. Econ. 87 (1973) 355–374.
[9] C. Meyer, Social finance and the commons paradigm: exploring how community-based innovations transform finance for the common good, Management Decision 58 (2020) 786–796.
[10] S.N. Cheung, Property rights in trade secrets, Econ. Inq. 20 (1982) 40–53.
[11] M. Demiray, S. Burnaz, Positioning of crowdfunding platforms: turkey as an emerging market case, Journal of Management Marketing and Logistics 6 (2019) 84–94.
[12] S. Kleinert, C. Volkmann, M. Grünhagen, Correction to: third-party signals in equity crowdfunding: the role of prior financing, Small Business Economics 54 (2020) 367–372.
[13] B.J. Cowden, S.L. Young, The copycat conundrum: the double-edged sword of crowdfunding, Bus. Horiz. 63 (2020) 541–551.
[14] L. Dezi, V. Cillo, A. Usai, P. Pisano, Equity crowdfunding in technology transfer strategies and licensing, Int. J. Technol. Manage. 78 (2018) 28–51.
[15] L. Valanciene, S. Jegeleviciute, Valuation of crowdfunding: benefits and drawbacks, Economics and Management 18 (2013) 39–48.
[16] H.-.G. Choe, S.-Y.T. Lee, Financial Footnote Analysis for Financial Ratio Predictions based on Text-Mining Techniques, Knowledge Management Research 21 (2020) 177–196.
[17] P. Craja, A. Kim, S. Lessmann, Deep learning for detecting financial statement fraud, Decision Support Systems 139 (2020), 113421.
[18] A. Pérez-Luño, J. Cambra, Listen to the market: do its complexity and signals make companies more innovative? Technovation 33 (2013) 180–192.
[19] M.J. Zhou, B. Lu, W.P. Fan, G.A. Wang, Project description and crowdfunding success: an exploratory study, Information Systems Frontiers 20 (2018) 259–274.
[20] J. Loewenstein, W. Ocasio, C. Jones, Vocabularies and vocabulary structure: a new approach linking categories, practices, and institutions, The Academy of Management Annals 6 (2012) 41–86.
[21] W. Wang, L. He, Y.J. Wu, M. Goh, Signaling persuasion in crowdfunding entrepreneurial narratives: the subjectivity vs objectivity debate, Comput. Hum. Behav. 114 (2021), 106576.
[22] W. Wang, W. Chen, K. Zhu, H. Wang, Emphasizing the entrepreneur or the idea? The impact of text content emphasis on investment decisions in crowdfunding, Decision Support Systems 136 (2020), 113341.
[23] H. Yuan, R.Y. Lau, W. Xu, The determinants of crowdfunding success: a semantic text analytics approach, Decision Support Systems 91 (2016) 67–76.
[24] W. Wang, L. Guo, Y.J. Wu, The merits of a sentiment analysis of antecedent comments for the prediction of online fundraising outcomes, Technological Forecasting and Social Change 174 (2022), 121070.
[25] M. Siering, J.-.A. Koch, A.V. Deokar, Detecting fraudulent behavior on crowdfunding platforms: the role of linguistic and content-based cues in static and dynamic contexts, Journal of Management Information Systems 33 (2016) 421–455.
[26] A. Parhankangas, M. Renko, Linguistic style and crowdfunding success among social and commercial entrepreneurs, Journal of Business Venturing 32 (2017) 215–236.
[27] Y. Koh, M. Lee, J. Kim, Y.Y. Yang, Successful restaurant crowdfunding: the role of linguistic style, International Journal of Contemporary Hospitality Management 32 (2020) 3051–3066.
[28] X. Yuan, H. Wang, How linguistic cues affect the motivation of capital-giving in crowdfunding: a self-determination theory perspective, New Review of Hypermedia and Multimedia 26 (2020) 1–23.
[29] F.P. Appio, D. Leone, F. Platania, F. Schiavone, Why are rewards not delivered on time in rewards-based crowdfunding campaigns? An empirical exploration, Technological Forecasting and Social Change 157 (2020), 120069.
[30] M.A. Stanko, D.H. Henard, Toward a better understanding of crowdfunding, openness and the consequences for innovation, Research Policy 46 (2017) 784–798.
[31] R. Jaziri, M. Miralam, Modelling the crowdfunding technology adoption among novice entrepreneurs: an extended TAM model, Entrepreneurship and Sustainability Issues 7 (2019) 353.
[32] Y.-.H. Tseng, C.-.J. Lin, Y.-.I. Lin, Text mining techniques for patent analysis, Information processing & management 43 (2007) 1216–1247.
[33] W. Drover, M.S. Wood, A.C. Corbett, Toward a cognitive view of signalling theory: individual attention and signal set interpretation, Journal of Management Studies 55 (2018) 209–231.
[34] M. Vaculík, A. Lorenz, N. Roijakkers, W. Vanhaverbeke, Pulling the Plug? Investigating Firm-Level Drivers of Innovation Project Termination, IEEE Trans. Eng. Manage. 66 (2019) 180–192.
[35] E. Mollick, A. Robb, Democratizing innovation and capital access: the role of crowdfunding, Calif. Manage. Rev. 58 (2016) 72–87.
[36] T. Wang, X. Liu, M. Kang, H. Zheng, Exploring the determinants of fundraisers' voluntary information disclosure on crowdfunding platforms, Online Information Review 42 (2018) 324–342.
[37] V. Babich, S. Marinesi, G. Tsoukalas, Does crowdfunding benefit entrepreneurs and venture capital investors? Manufacturing & Service Operations Management 23 (2021) 508–524.
[38] A.B. Radnejad, H. Vredenburg, J. Woiceshyn, Meta-organizing for open innovation under environmental and social pressures in the oil industry, Technovation 66 (2017) 14–27.
[39] M.-.Y. Chen, C.-.H. Liao, E.D. Lughofer, E. Egrioglu, Informetrics on social network mining: research, policy and practice challenges, Library Hi Tech 38 (2020) 273–275.
[40] B.F. Klimova, Using corpus linguistics in the development of writing, Procedia-Social and Behavioral Sciences 141 (2014) 124–128.
[41] G.R. Klare, Assessing readability, Reading Research Quarterly (1974) 62–102.
[42] N.T. Demoulin, K. Coussement, Acceptance of text-mining systems: the signaling role of information quality, Information & Management 57 (2020), 103120.
[43] L. Zhou, A. Fu, Y. Mu, H. Wang, S. Yu, Y. Sun, Multicopy provable data possession scheme supporting data dynamics for cloud-based Electronic Medical Record system, Information Sciences 545 (2021) 254–276.
[44] Y. Chen, H. Zhang, R. Liu, Z. Ye, J. Lin, Experimental explorations on short text topic mining between LDA and NMF based schemes, Knowledge-Based Systems 163 (2019) 1–13.
[45] H. Chalu, The effect of IFMIS adoption on financial reporting quality in Tanzanian local governments, Business Management Review 22 (2020) 1–31.
[46] P. Crosetto, T. Regner, It's never too late: funding dynamics and self pledges in reward-based crowdfunding, Research Policy 47 (2018) 1463–1477.
[47] N. Korfiatis, E. GarcíA-Bariocanal, S. SáNchez-Alonso, Evaluating content quality and helpfulness of online product reviews: the interplay of review helpfulness vs. review content, Electronic Commerce Research and Applications 11 (2012) 205–217.
[48] J.D. Gregorio, J.W. Lee, Education and income inequality: new evidence from cross-country data, Review of income and wealth 48 (2002) 395–416.
[49] T.S. Campbell, J.K. Dietrich, The determinants of default on insured conventional residential mortgage loans, The Journal of Finance 38 (1983) 1569–1581.
[50] H.-.H. Hsu, Y.-.S. Chen, C.-.J. Lin, T.-.W. Pai, Patient Oriented Readability Assessment for Heart Disease Healthcare Documents, International Journal of Data Warehousing and Mining (IJDWM) 16 (2020) 63–72.
[51] H. Si, S. Jiang, Y. Fang, U. Muhammad, Can Readability of Loan Description Affect Loan Success Rate and Loan Cost?: a Textual Analysis of P2P Loan Description, Engineering Economics 31 (2020) 302–313.
[52] A.H. Anglin, M.T. Wolfe, J.C. Short, A.F. McKenny, R.J. Pidduck, Narcissistic rhetoric and crowdfunding performance: a social role theory perspective, Journal of Business Venturing 33 (2018) 780–812.
[53] Q. Liu, Q. Du, Y. Hong, W. Fan, S. Wu, User idea implementation in open innovation communities: evidence from a new product development crowdsourcing community, Information Systems Journal (2020).
[54] J.-.M. Kim, D.M. Im, S. Jun, Factor analysis and structural equation model for patent analysis: a case study of Apple's technology, Technology Analysis & Strategic Management 29 (2017) 717–734.
[55] L. Li, T.-.T. Goh, D. Jin, How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis, Neural Computing and Applications 32 (2020) 4387–4415.
[56] A.F. McKenny, T.H. Allison, D.J. Ketchen Jr, J.C. Short, R.D. Ireland, How should crowdfunding research evolve? A survey of the entrepreneurship theory and practice editorial board, Entrepreneurship Theory and Practice 41 (2017) 291–304.
[57] D. Gedda, B. Nilsson, Z. Såthén, K.S. Søilen, Crowdfunding: finding the optimal platform for funders and entrepreneurs, Technology Innovation Management Review 6 (2016) 31–40.
[58] J.-.F. Soublière, J. Gehman, The legitimacy threshold revisited: how prior successes and failures spill over to other endeavors on Kickstarter, Acad. Manag. J. 63 (2020) 472–502.

[59] C.W. Ricci, Annual report readability: the case of small-cap companies, International Research Journal of Applied Finance 8 (2017) 712–718.
[60] P.-.J. Lee, Y.-.H. Hu, K.-.T. Lu, Assessing the helpfulness of online hotel reviews: a classification-based approach, Telematics and Informatics 35 (2018) 436–445.
[61] A.E. Eltorai, S.S. Naqvi, S. Ghanian, C.P. Eberson, A.P.C. Weiss, C.T. Born, A. H. Daniels, Readability of invasive procedure consent forms, Clinical and Translational Science 8 (2015) 830–833.
[62] G. Spache, A new readability formula for primary-grade reading materials, The Elementary School Journal 53 (1953) 410–413.
[63] G.H. McLaughlin, SMOG grading-a new readability formula, Journal of Reading 12 (1969) 639–646.
[64] Y. Guan, K.A. Maloney, D.L. Roter, T.I. Pollin, Evaluation of the informational content, readability and comprehensibility of online health information on monogenic diabetes, J. Genet. Couns. 27 (2018) 608–615.
[65] R. Gunning, The fog index after twenty years, Journal of Business Communication 6 (1969) 3–13.
[66] E. Fry, A readability formula that saves time, Journal of Reading 11 (1968) 513–578.
[67] S.A. Crossley, D.B. Allen, D.S. McNamara, Text readability and intuitive simplification: a comparison of readability formulas, Reading in a Foreign Language 23 (2011) 84–101.
[68] W. Wang, K. Zhu, H. Wang, Y.-C.J. Wu, The impact of sentiment orientations on successful crowdfunding campaigns through text analytics, IET Software 11 (2017) 229–238.
[69] A.C. Poisson, I.M. McCullough, K.S. Cheruvelil, K.C. Elliott, J.A. Latimore, P. A. Soranno, Quantifying the contribution of citizen science to broad-scale ecological databases, Frontiers in Ecology and the Environment 18 (2020) 19–26.
[70] H. Sauermann, K. Vohland, V. Antoniou, B. Balázs, C. Göbel, K. Karatzas, P. Mooney, J. Perelló, M. Ponti, R. Samson, Citizen science and sustainability transitions, Research Policy 49 (2020), 103978.
[71] Y. Li, F. Liu, W. Fan, E.T. Lim, Y. Liu, Exploring the impact of initial herd on overfunding in equity crowdfunding, Information & Management (2020), 103269.
[72] Y. Hong, Y. Hu, G. Burtch, Embeddedness, pro-sociality, and social influence: evidence from online crowdfunding, MIS Quarterly 42 (2018) 1211–1224.
[73] J. Cha, Crowdfunding for video games: factors that influence the success of and capital pledged for campaigns, International Journal on Media Management 19 (2017) 240–259.
[74] M. Kang, Y. Gao, T. Wang, H. Zheng, Understanding the determinants of funders' investment intentions on crowdfunding platforms: a trust-based perspective, Industrial Management & Data Systems 116 (2016) 1800–1819.
[75] I. Kocoglu, G.S. Lynn, Y. Jung, P.G. Dominick, Z.H. Aronson, P. Burke, Actions speak louder than words: how team trust and commitment enhance team action listening and team success, Management Decision 58 (2020) 465–494.
[76] C. Ai, E.C. Norton, Interaction terms in logit and probit models, Economics letters 80 (2003) 123–129.
[77] T. Tammi, J. Saastamoinen, H. Reijonen, Public procurement as a vehicle of innovation–What does the inverted-U relationship between competition and innovativeness tell us? Technological Forecasting and Social Change 153 (2020), 119922.
[78] H. Pakhchanian, R. Raiker, C. Geist, Readability Analysis of Dacryocystorhinostomy Online Patient Information, Invest. Ophthalmol. Vis. Sci. 61 (2020), 2941-2941.
[79] H. Cao, S. Bernard, R. Sabourin, L. Heutte, Random forest dissimilarity based multi-view learning for Radiomics application, Pattern Recognit. 88 (2019) 185–197.
[80] T. Vanacker, D.P. Forbes, M. Knockaert, S. Manigart, Signal strength, media attention, and resource mobilization: evidence from new private equity firms, Acad. Manag. J. 63 (2020) 1082–1105.
[81] F.M. Saluzzo, I. Alegre, Supporting entrepreneurs: the role of third-party endorsement in crowdfunding platforms, Technological Forecasting and Social Change 162 (2021), 120402.
[82] P.R. Murphy, L. Purda, D. Skillicorn, Can fraudulent cues be transmitted by innocent participants? Journal of Behavioral Finance 19 (2018) 1–15.
[83] F. Di Pietro, M.L. Bogers, A. Prencipe, Organisational barriers and bridges to crowd openness in equity crowdfunding, Technological Forecasting and Social Change 162 (2021), 120388.
[84] M.L. Martens, J.E. Jennings, P.D. Jennings, Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition, Acad. Manag. J. 50 (2007) 1107–1132.

Wei Wang, PhD, is an associate professor at the Department of Information Management, College of Business Administration, at Huaqiao University. He holds a PhD from Tongji University in Shanghai, a master's and a bachelor degree from Huaqiao University in Quanzhou, China. His research interests include crowd funding, sentiment analysis, and text analysis. His work has appeared in academic journals including *Decision Support Systems, Computers in Human Behavior, Behaviour & Information Technology, Industrial Management & Data Systems*, Journal of Experimental & Theoretical Artificial Intelligence, Management World (China), *Journal of Management Sciences in China, Systems Engineering—Theory & Practice*, among others.

Yuting Xu is a postgraduate student at the Department of Business Administration, at Huaqiao University. Her research interests include crowd funding and business management.

Yenchun Jim Wu (Ph.D, University of Michigan) is Distinguished Professor of Graduate Institute of Global Business and Strategy, College of Management at National Taiwan Normal University, Taiwan. His papers have appeared in *Academy of Management Learning and Education, MIT Sloan Management Review, IEEE Transactions on Engineering Management, Information & Management, Computers in Human Behavior, Computers & Education, International Journal of Operations & Production Management, Supply Chain Management, Management Decision, Journal of the American Society for Information Science and Technology, International Journal of Physical Distribution & Logistics Management, International Journal of Logistics Management, European Journal of Operational Research, International Journal of Production Economics, Transportation Research Part A,D & E, and Technological Forecasting and Social Change*. His research interests include supply chain management, technology management, and innovation and entrepreneurship, and management education.

Mark Goh holds a Ph.D. from the University of Adelaide. In the National University of Singapore, he holds the appointments of Director (Industry Research) at the Logistics Institute-Asia Pacific, a joint venture with Georgia Tech, USA, Principal Researcher at the centre for Transportation Research, and was a Program Director of the Penn-State NUS Logistics Management Program. His current research interest is on buyer-seller relationships, performance measurement, and supply chain strategy. With over 400 technical papers in internationally refereed journals and conferences, and won research grant money in excess of SGD 10 million, some of his articles on supply chain management have appeared in the Journal of Purchasing and Materials Management, Industrial Marketing Management, European Journal of Purchasing and Supply Chain Management, IIE Transactions, Naval Research Logistics, Physical Distribution and Logistics Management, Production and Operations Management, EJOR, Supply Chain Management Journal, Industrial Organisations, Journal of Global Optimization, Journal of Asian Business, SIAM Journal of Optimization, Transportation Research Parts A, D & E, MIS Quarterly Executive, JOTA, International Journal of Production Research, Decision Sciences, IEEE Trans in Engineering Management, Transportation Science, PLOSOne, JAIS, and Logistics Information Management.