

Understanding the Effect of Task Descriptions on User Participation in Crowdsourcing Contests: A Linguistic Style Perspective

Shuang Wu
School of Information
Central University of
Finance and Economics
wushuang96@163.com

Qian Liu
Central University of
Finance and Economics
liuqianlf@cufe.edu.cn

Baowen Sun
Central University of
Finance and Economics
sunbaowen@263.net

Xin Zhao
School of Economics
and Management
Xi'an of Technology
zhaoxin_zzz@163.com

Abstract

Many employers are struggling with how to deliver attractive tasks on crowdsourcing platforms, where users can be effectively integrated into a company's tasks. In this study, the linguistic style of crowdsourcing task descriptions is investigated, and an analysis is conducted on how such linguistic styles are related to a task description's success in attracting participants. Based on uncertainty reduction theory as well as source credibility theory, an empirical analysis of 2,014 designing contests demonstrates that certain linguistic styles will reduce the uncertainty perceived by crowdsourcing solvers and increase employers' credibility, generating positive effects on participation. It is also found that these observed effects are moderated by the magnitude of the rewards offered for completing crowdsourcing tasks. The results of this study inform the theories concerned on crowdsourcing participation, linguistics, as well as psychological processes, while offering the industry insight on how to describe their own crowdsourcing tasks better.

1. Introduction

Online innovation contests, conducted through websites such as Innocentive, are becoming an increasingly important source for many companies and institutions to have tasks completed by relying on the capabilities of an entire crowd instead of a single employee or machine. Utilizing the experience and effort of large number of people via the Internet is commonly referred to as crowdsourcing, and it has been used effectively in a number of variable applications([4]; [9]). By crowdsourcing, widespread communities of people can be effectively integrated into a company's tasks([4]; [31]).

A central component of any crowdsourcing task is the task's description, in which a given employer

presents the task to be solved. These descriptions are typically detailed introductions of the given company's published task. Potential solvers who interested in the task will read this content, gain an understanding of the company's requirements, and decide whether they shall participate or not. Thus, it can be seen that the manner in which employers deliver their requirement to their prospective solvers is critical in the effective crowdsourcing of tasks. In this study, we focus on the linguistic style of crowdsourcing tasks and how such a style relates to the success in raising participations.

Past research has gradually found that task descriptions indeed have an impact on participation ([3]; [24]). However, empirical studies on the topic tend to focus primarily on the length of a task title, the length of its description, and the utility of manual annotation to analyze the level of detail ([23]; [33]). It must be noted that some studies have presented conflicting conclusions, with a few studies finding no significant effects from task descriptions ([3]; [33]; [32]). Missing from this body of research is evidence on how the style of linguistic factors contributes to the success of crowdsourcing contests. In this study, the role of language is examined closely in the context of crowdsourcing participation. Based on uncertainty reduction theory and source credibility theory, some specific linguistic styles such as message concreteness, cognitive complexity are introduced into the study's prediction model to investigate. Taken from the research literature on psycholinguistics and in the context of interpersonal persuasion, these variables may well be effective in reducing user uncertainty and increasing employer credibility ([12]; [29]; [30]), then generating positive effects on participation. It is also found that the role played by these linguistic styles in task descriptions is related to the amount of reward offered for completing a given task.

This study contributes to the research literature on the topic in three major ways. First, it is one of the first studies to explain the role task descriptions play in crowdsourcing participation from the perspective of

linguistic style. Second, this study contributes to the field in its application of research on linguistic style into a new area, namely crowdsourcing contests. Third, this study takes the difference between users' initial attention and actual participation into consideration, finding the role linguistic style plays in transforming prospective solvers to actual participants.

Beyond academic theory, these findings also have important implications for employers that offer tasks on various crowdsourcing platforms. Although most people are not explicitly aware of the subtle differences in linguistic style that present themselves in daily life, the findings of this study show that by writing task descriptions carefully, crowdsourcing employers can present their demands, intentions, and expectations both adequately and reliably to elicit a greater response from the willing and able crowd.

2. Literature Review

2.1. Crowdsourcing Contest

Crowdsourcing contests have become a particularly popular and effective way for companies to integrate the variable capabilities of large groups of people into specific tasks they need to complete [4]. Employers will offer tasks through an online network platform, and then prospective solvers will choose whether to participate in the task based on their own considerations. If these solvers choose to tackle a given task, they will use their professional knowledge and skills to provide the company with creative and efficient solutions. The solvers who submit the optimal solution will win the crowdsourcing contest and receive a reward for their time and effort [28]. Through such crowdsourcing competitions, many enterprises will not only obtain creative solutions to their problems, but also save time, manpower, and economic costs through the medium of crowdsourcing([4]; [31]).

The accelerating growth of crowdsourcing contests in recent years has garnered considerable attention in academic circles. A central question among these researchers is the reason why participants choose to share their ideas and how one can design a task which attracts greater numbers of prospective solvers. Past research conducted in this vein gradually concluded that task descriptions have a significant impact on user participation ([3]; [24]). One research has proposed that employers should describe their crowdsourcing task clearly from the moment the crowdsourcing contest begins in order to attract higher numbers of potential solvers [24]. Other research has found that in order to improve the efficiency of a crowdsourcing competition and to also improve innovation performance, employers

should disclose private information related to the crowdsourcing task itself while at the same time control the flow of key information so as not to leak such key information to their competitors [3]. Switching cost theory also suggests that a contest with higher learning costs will present higher switching costs, and as such, a contest with longer description will attract fewer prospective solvers [33].

Contemporary research on the topic suggests that task descriptions affect the successful crowdsourcing of tasks, however, past empirical studies have focused primarily on questionnaire research or have worked through manually annotated regression models based on the length of the task title, the length of its description, or the description's level of detail ([23]; [33]). The findings of some studies have even been found to have conflicting conclusions ([24]; [33]). Moreover, other studies have been unable to find a significant effect behind the task descriptions of crowdsourcing contests [32]. Missing from this body of research is evidence on how the linguistic styles found in task descriptions factor into crowdsourcing success. This study seeks to address this gap in the research.

2.2. A Note on Linguistic Style

Linguistic style provides us with important clues on how people process and interpret information, and how they respond to the statements that contain psychological significance. Past research has shown that the use of functional words, such as personal pronouns, can reveal people's implicit intentions ([19]; [20]). For example, people who are experiencing physical or emotional pain will tend to pay more attention to themselves and their own situation, and therefore will use singular first-person pronouns more often [19]. The use of pronouns is also closely related to the quality of intimacy. When married couples were asked to evaluate their marriages with interviewers, the more the subjects used the pronoun "we", the better their marital status was found to be [20].

Content words are likewise very significant in human interaction [29]. The use of causal words (e.g., "making" and "becoming") and insightful words (e.g., "understanding" and "realizing") can reflect the process of reevaluating events [30]. When writing of their traumatic experiences of individual, individuals will use more causal words and insight words, because doing so activates the processing of the event in an individual [26]. When people are uncertain about a specific topic, they will lean toward using more tentative words (e.g., "presumably" and "almost") [15]. Studies have also found that for the lies published in instant messaging, the amount of words used was comparatively greater

when lying, while the number of first-person pronouns used was found to be reduced when similarly lying [8].

Utilizing words as implicit behavioral indicators can work to avoid many of the reaction biases or social approbation problems that are commonly faced by self-responsive questionnaires [17]. Analyzing the use of such words can provide cues for psychological analyses and can also serve as an indicator for behavior in the study of psychological mechanisms [17]. The linguistic style of task descriptions may imply more meaningful information that works to reduce uncertainty and elucidate credibility in crowdsourcing. In the following section, the form of linguistic style that takes effect in the crowdsourcing contests is explained.

3. Theories and Hypotheses

3.1. Theories

Uncertainty reduction theory (URT), a theory with its roots in system science and psychology, was first presented by Berger and Calabrese in 1975 [2]. “Uncertainty” which is also often referred to as “unstable feelings”, denotes an individual’s mental state of uncertainty when he or she is unable to make specific and clear identification and evaluation of his or her own psychology and behavior [2]. Generally speaking, uncertainty exhibits a dynamic effect. It encourages an individual to change his or her current situation by seeking information to acquire knowledge in order to achieve a psychological sense of certainty. Crowdsourcing contests represent situations that are filled with uncertainty. First, the employers on online crowdsourcing platforms are typically complete strangers to their potential solvers, and they often cannot contact each other in a commonly timely manner. Secondly, many of the participants in a given crowdsourcing project may not in fact be able to provide work that meets the requirements of a crowdsourcing employer. Thirdly, the employer will select a successful bidder from the work provided by many different participants. This means that the effort of many of the participants does not necessarily pay off at all. It follows, then, that participants are oftentimes forced to rely on the linguistic clues they find in a crowdsourcing project’s task description to help them understand the needs of the tasks’ employer, thereby increasing their chances of winning the contest.

Source credibility refers to the authenticity, credibility, and integrity of the source of information, as perceived by such users [18]. Past studies on the topic have found that source credibility has a direct impact on the attitude of a user. Specifically, a user’s high confidence in source information increases information

credibility [5]. In the crowdsourcing context, past studies have shown that participants often seek an employer with a higher brand strength and higher levels of credibility [32]. This is because participants often expect their employers to be highly trustworthy, professional, as well as objective, and expect that such employers will carefully select winning bidders from the numerous submissions made for a project then pay for it. Solvers can detect clues on source credibility in the linguistic style utilized in the task description of a given crowdsourcing project.

3.2. Hypotheses

Building on past research on psycholinguistics and communication, it is proposed in this study that solvers seek and utilize linguistic cues in order to reduce uncertainty and infer the credibility of their potential employers, which in turn impacts the rate of solver participation in a crowdsourcing project. In brief, a task description needs to exhibit some specific characteristics like message concreteness, cognitive complexity, psychological distance, the preciseness of language and intellectual property declaration. These characteristics are elaborated upon each, below.

Message Concreteness: The term concreteness denotes a kind of linguistic style which is closely related to uncertainty reduction as reinforced in literature ([29]; [30]). Concrete words function as contextual and detailed representations that allow readers to process information faster and more easily ([13]; [14]; [27]). In the crowdsourcing context, a concrete description can provide more specific information that better signals an employer’s requirements, and hence can function more effectively to reduce uncertainty and develop solver confidence in completing work that is in the end more consistent with the original intention of the employer.

Two linguistic cues have been identified to describe the concreteness of a text. They consist of both the statement’s articles (i.e. “a”, “an”, and “the”) [26] as well as its quantifiers (e.g., “lots”, “a few”, and “a lot”) [12]. Research in the field has already proven the important role concrete words play in certain online contexts, such as the case for peer-to-peer lending sites [12] and online lie detecting [30]. Concrete words are very likely associated with crowdsourcing participation as well. This is particularly the case when one considers the fact that the requirements of a potential crowdsourcing employer are complex, and therefore present difficulties for achieving solver’s understanding at the outset. Hence, the following hypothesis is presented.

H1(a): Crowdsourcing task descriptions that contain more concrete language (by using more articles) will exhibit increased participation.

H1(b): Crowdsourcing task descriptions that contain more concrete language (by using more quantifiers) will exhibit increased participation.

Cognitive Complexity: In their drive to reduce uncertainty, solvers will value the depth of thinking and reasoning that is presented in the messages employers will transmit to state their requirements. This depth of thinking, as well as the details of an employer's preferences, are manifested in cognitive complexity [21]. Past studies note that negation words (e.g., "no", "not", and "never") can effectively demonstrate complex thinking on behalf of the signaler, because they are more specific and precise [8]. Moreover, negation words differentiate between what belongs to a category and what does not, in itself a cognitively complex task ([1]; [16]). Negation words help solvers by making it clearer what an employer truly wants and what an employer does not want, thereby reducing uncertainty. Hence, the following hypothesis is presented.

H2: Crowdsourcing task descriptions that contain more markers of cognitive complexity (i.e. more negation words) will exhibit increased participation.

Psychological Distancing: Psychological distance is a kind of social psychology term that expresses the degree of subjective feeling that an individual is intimate, accommodating, or difficult to get along with another body or group [30]. In a crowdsourcing context, it is the employer that chooses the best solution from a large pool of solver contribution and offers the reward [33]. Coming from this bargaining position, an employer need not to present a lower psychological distancing. On the contrary, employers will be expected to conduct themselves in an objective, detached, and unbiased manner. This psychological distancing, which can be thought of as a separation between the employers and its potential crowdsourcing solvers, has been shown to manifest itself linguistically in a decreased rate of first-person pronoun usage [6]. Recall that personal pronouns work to reveal people's intentions. Consistent with this, the people who pay more attention to themselves are those who exhibit greater rates of first-person pronoun usage ([6]; [19]).

People also lower their psychological distance by expressing their negative emotions [29]. This is especially the case when it comes to anxiety. As such, another linguistic mark of psychological distancing is the utilization words that are categorized as less anxiety focused [29]. Indeed, employers should not display signs of anxiety when writing out their crowdsourcing task descriptions, as anxiety represents an emotional state that is antithetical to objective thinking. Hence, the following hypothesis is presented.

H3(a): Crowdsourcing task descriptions which are low in psychological distancing (by using fewer first-person pronouns) will exhibit increased participation.

H3(b): Crowdsourcing task descriptions which are low in psychological distancing (by using fewer anxiety related words) will exhibit increased participation.

Precise Language: Early research on psycholinguistics showed that if a text is too rich in the vocabulary, the linguistic richness affects the rigor of the information provided. As a consequence of this, a target audience will assume more potential deception on the part of the writer ([7]; [11]). The preciseness of the language used in a text is inherently contrary to lexical diversity, however [10]. For instance, an employer may use the synonyms "design", "work", "production", "result", or "outcome" to describe the requirements it may have for a specific task. This will make the description lacking in preciseness. As an example, taken from the opposite perspective, an employer that uses "design" consistently throughout the entirety of the text will offer a text with precise language. One can expect that using language in the task description for a crowdsourcing project that is imprecise will work to alienate potential solvers by creating the appearance of bias, unprofessionalness, and lacking credibility. Hence, the following hypothesis is presented.

H4: Crowdsourcing task descriptions that use precise language will exhibit increased participation.

Intellectual Property Declaration: An Intellectual Property Declaration (IP) declaration describes any problems an employer may run into regarding intellectual property rights. On most crowdsourcing platforms, and specifically for design tasks, intellectual property declarations will require the winning solver's intellectual property to be transferred wholly to the rewarding employer. Though the website has already made the above statement, the restatement in the task description helps to reflect the employer's perceived objectivity, stringency, and credibility, which gives the solvers greater confidence in an employer's ability to select winning bidders carefully and then pay the promised reward. This, in turn, will work to increase crowdsourcing participation. Hence, the following hypothesis is presented.

H5: Crowdsourcing task descriptions which have an IP declaration will exhibit increased participation.

The Moderation Effect of Task Reward: For this study, task reward is also considered as a moderator in these linguistic styles. The greatest motivation for participants to involve themselves in a crowdsourcing task is in their ability to obtain a reward, most typically a monetary one ([22]; [25]). In line with the risk premium theory in economics, the potential solvers that are involved in crowdsourcing tasks are inherently risk averse. They pay many costs, such as investments in time and energy, to complete these crowdsourcing tasks, and the benefits of doing so are oftentimes quite uncertain. Therefore, in order to encourage participants

to ignore such potential risks, crowdsourcing employers will often increase their task rewards, so that the expected benefits for these solvers will better offset the potential costs invested.

For this reason, the amount of the reward offered for a given task serves as a critical regulatory variable for gauging solver uncertainty and source credibility. As the magnitude of the reward increases, the risk tolerance of solvers will increase, and they will pay less attention to a task description’s language style in turn. This means that, when the reward offered is large enough, and regardless of whether the role of these linguistic cues is positive or negative, the psychological influence upon participants will be weakened. Hence, the following hypotheses are presented.

H6(a): The amount of the reward offered will exhibit a negative effect on the role of concreteness (using more articles) in crowdsourcing task descriptions.

H6(b): The amount of the reward offered will exhibit a negative effect on the role of concreteness (using more quantifiers) in crowdsourcing task descriptions.

H7: The amount of the reward offered will exhibit a negative effect on the role of cognitive complexity in crowdsourcing task descriptions.

H8(a): The amount of the reward offered will exhibit a negative effect on the role of psychological distancing (using fewer first-person pronouns) in crowdsourcing task descriptions.

H8(b): The amount of the reward offered will exhibit a negative effect on the role of psychological distancing (using fewer anxiety related words) in crowdsourcing task descriptions.

H9: The amount of the reward offered will exhibit a negative effect on the role of using precise language in crowdsourcing task descriptions.

H10: The amount of the reward offered will exhibit a negative effect on the role of having an IP declaration in crowdsourcing task descriptions.

All derived hypotheses are collected and shown in Figure 1.

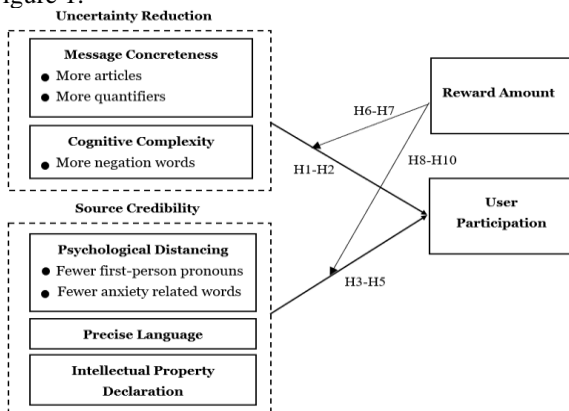


Figure 1. Hypotheses of Linguistic Styles on User Participation

4. Method

4.1. Sample

This study investigates a large-scale dataset of real world crowdsourcing contests conducted online. The data were collected from TaskCN.com, one of the major crowdsourcing platforms of China. Within just six months of its establishment in early 2006, the website has occupied about 60% of the domestic crowdsourcing market of China and has become a leader in the industry. By the end of 2017, the site had had over 3.6 million registered solvers and had hosted over 35,000 contests. In this platform, employers can start a contest with award deposit. Solvers first browse the task list, after getting information like the task title, task type, reward, employer’s credit score, they select the task that interests them and click into the task details page. There, solvers can read the task description information and decide whether to participate or not.

The study’s sample data were collected over three years, from January 2015 to January 2018. In total, approximately 3,700 single-winner contest projects are included. For the following reasons, only the category of design tasks was selected for analysis. Firstly, with the development of the website, the proportion of design tasks has increased in recent years. Over 80% of the newly released tasks in the past three years are design tasks, which have gained considerable attention and large number of participants. Secondly, the design task is a kind of task that requires more creativity. Employers must fully express their needs while at the same time encourage participants to maintain their unique thinking. As a consequence, the writing of task descriptions is of vital importance. Thirdly, the design tasks, which have a certain degree of specialty and difficulty, require participants to make more efforts, so they must read the task descriptions more carefully in order to evaluate the effort and the reward. In short, for design tasks, both employers and solvers pay higher attention on the content of task descriptions. As a result, other types of tasks (e.g., writing, coding, and translating) were removed from the sample. After eliminating some tasks which cannot crawl to the task description, the final total for the study’s sample data is 2,014 contests.

4.2. Variable Measurements

Dependent Variable: The study’s dependent variable, the participation of a crowdsourcing contest, can be measured specifically as the number of solvers which is directly available on the crowdsourcing platform.

Table 1. Descriptive Statistics of Key Variables

Variable	Mean	Std Dev.	Max	Min
Reward Amount	478.98	493.403	5000	100
Quantifier	1.63	1.665	17	0
Article	0.02	0.147	3	0
Negation	0.23	0.393	6	0
IP Declaration	0.35	0.476	1	0
First-pronoun	0.31	0.942	12	0
Anxiety	0.02	0.151	3	0
Precise	-74.56	15.306	-28	-100
Followers	4303.48	2288.929	43571	608
Number of Solvers	24.50	17.929	152	1
Observation Numbers	2014			

Independent Variables: The study’s independent variables (e.g., message concreteness, cognitive complexity, and psychological distancing) can be analyzed using LIWC 2015 software. Linguistic Inquiry and Word Count (LIWC) is a type of natural language processing that works to quantify the content of a given text and calculate the different types of words used therein. LIWC consists primarily of two parts, the program and its dictionary. The LIWC dictionary defines category names and word lists for word attribution. The LIWC program itself compares the words in a sample text with this dictionary by importing a sample dictionary and text, producing word frequency results for the various types of words found. LIWC2015 has a new function for Chinese language processing. For this study, each task description is converted to a text file and assigned an identification number. Then code was written to perform Chinese word segmentation on each individual text. Thereafter, the LIWC program processed each task description separately, producing an output that indicates the word frequency for each category.

For this study’s analysis, the effects of individual language cues (e.g., pronouns and anxiety words) are reported. Rather than combining these language cues into composite indicators (e.g., psychological distances), this is done so that this study’s research results can be directly compared to the existing scholastic literature on the topic. Additionally, research shows that linguistic correlates to the same psychological construct sometimes operate in opposite directions than hypothesized across different contexts [29]. Given that the study’s composite index may obscure these operations by averaging a variety of language categories, simply reporting the impact of individual language cues presents a more reliable methodology.

However, the two variables, precise language and intellectual property declarations, cannot be measured by using LIWC. Precise language can be measured as the negative (or opposite) of the type-token ratio (TTR), which is sometimes also referred to as the inferiority ratio. The term inferiority ratio refers to the ratio of the classifiers to impersonators that appear in a sample text. The type (or character) refers to all non-repeating words in a sample text. The token refers to all the words in the entire text, including those that are repeatedly used. A higher value for TTR will indicate cautiousness and deft manipulation on the part of the communicator (here, the crowdsourcing employer), while a lower value for TTR is associated with a more relaxed and transparent communication style ([10]; [11]). Wordsmith software was utilized in this study to obtain the TTR of each task description sample text, taken as a negative value more specifically.

The IP declarations in the sample data were operationalized as dummy variables that indicate whether a given entry has an IP declaration or not. The parsing program was coded to process the text information contained in each task description, specifically analyzing whether the words “intellectual property”, “property right”, and “copyright” were

Table 2. Correlation Matrix of Key Variables

	1	2	3	4	5	6	7	8	9	10
1 Number of Solvers	1									
2 Reward	0.456***	1								
3 Followers	0.573***	0.45***	1							
4 Quantifier	-0.010	0.065**	0.039*	1						
5 Article	0.024	0.025	0.058**	0.009	1					
6 Negation	0.121	0.108	0.082	0.269	-0.015	1				
7 IP Declaration	***	***	***	***	***	0.680	1			
8 First-pronoun	***	***	***	***	***	***	***	1		
9 Anxiety	-0.068***	-0.017	-0.039*	-0.006	0.04*	-0.063**	-0.066***	0.049*	1	
10 Precise	-0.004	0.009	0.006	0.024	0.000	-0.002	0.012	0.034	-0.034	1
	0.008	-0.005	-0.003	0.018	0.027	0.032	0.023	0.034	-0.034	1

* p < 0.05; **p < 0.01; ***p < 0.001.

included. If these aforementioned terms were found, the variable was marked as 1. If they were not found, the variable was marked as 0.

Control Variables: Following earlier studies on comparable online contests, this study controls several factors that have been found to be associated with crowdsourcing success. These control variables are the market maturity (task ID), task reward amount, task duration (in minutes), task title length (in the number of Chinese characters), task description length (in the number of Chinese characters), the number of followers, an employer's credit score, registration time, cumulative number of posted tasks, and whether the employer used an email or real name authentication.

The number of followers represents the number of users who click to access the task details page. Only when a potential solver clicks on the details page can he or she see the task description text and then decide whether to actually participate or not. This is a very important control variable, after introducing the variable, the task description can really play a role. And the model's R square is significantly increased.

The descriptive analysis and correlation matrix are presented in Table 1 and Table 2, above.

5. Results and Analysis

For testing this study's hypotheses, three regression models were conducted and have passed the test of statistical assumptions of multiple regression. A base model that included only the control variables was run first (see Table 3, Model 1). In line with the past research conducted on these variables, the market maturity, reward amount, number of followers, task duration, and task title length were all found to be associated with crowdsourcing participation ([32]; [33]).

For Model 2, it is hypothesized that the linguistic cues that pertain to concrete language (i.e. articles and quantifiers), psychological distancing (i.e. personal pronouns and anxiety-related words), cognitive complexity (i.e. negations), precise language (the opposite of TTR), as well as IP declarations will be associated with solver participation rates. It was found that the model fits the data well, that the p-value is less than 0.001, and the model explained 48.9% of the variance in the dependent measure ($R = 0.702$, $R^2 = 0.493$, $R^2_{adj} = 0.488$). Negations, quantifiers, personal

Table 3. Regression Results

	Model 1	Model 2	Model 3
1 Constant	-130.913** (46.222)	-132.945** (45.758)	-130.870** (45.656)
2 Task ID	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
3 Reward Amount	0.007*** (0.001)	0.008*** (0.001)	0.008** (0.003)
4 Followers	0.005*** (0.000)	0.005*** (0.000)	0.005*** (0.000)
5 Duration	0.000* (0.000)	0.000* (0.000)	0.000* (0.000)
6 The Length of Description	-0.001 (0.001)	-0.007*** (0.001)	-0.006*** (0.001)
7 The Length of Title	-0.357*** (0.064)	-0.336*** (0.064)	-0.340*** (0.064)
8 Credit Score	0.014 (0.018)	0.014 (0.018)	0.016 (0.018)
9 Employers' Registration Time	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
10 Employers' Cumulative Number of Releases	-0.110 (0.061)	-0.132* (0.061)	-0.125* (0.061)
11 Email Authentication	-0.414 (0.658)	-0.025 (0.652)	-0.144 (0.650)
12 Real-name Authentication.	0.937 (0.879)	0.832 (0.869)	0.871 (0.866)
13 Quantifier		-0.622*** (0.183)	-0.502* (0.250)
14 Article		-0.390 (1.954)	-0.717 (3.427)
15 Negation		3.246*** (1.002)	-1.025 (1.556)
16 IP Declaration		4.112*** (1.139)	6.831*** (1.411)
17 First-pronoun		-0.826** (0.306)	-0.029 (0.414)
18 Anxiety		0.828 (1.905)	2.724 (2.972)
19 Precise		0.016 (0.019)	0.025 (0.026)
20 Quantifier*Reward Amount			0.000 (0.000)
21 Article*Reward Amount			0.001 (0.006)
22 Negation*Reward Amount			0.009*** (0.003)
23 IP Declaration*Reward Amount			-0.006*** (0.002)
24 First-pronoun*Reward Amount			-0.002** (0.001)
25 Anxiety*Reward Amount			-0.005 (0.007)
26 Precise*Reward Amount			-0.000 (0.000)
R ²	0.478	0.493	0.499
R ² _{adj}	0.475	0.488	0.493
SigF _{change}	0.000***	0.000***	0.001***
N	2014		

a * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; for the two-tailed test. b the values in parentheses denote standard deviations.

pronouns, and IP declarations all achieved statistical significance. However, precise language, articles, and anxiety were found not to be significant. Additionally, the number of quantifiers was found to have the opposite effect from what was predicted. Specifically, as the number of quantifiers decreased, higher rates of participation were exhibited (see Table 2). Although a higher number of observed quantifiers has been shown to be a concrete measure for task description specificity in other studies, this last finding may represent a key differentiation between online crowdsourcing task descriptions versus the task descriptions found for other online media. For instance, some quantifiers (e.g., “few,” “many,” “lots”) embody a certain amount of imprecision, where exact quantities are not explicitly specified. Online crowdsourcing solvers may desire more precise statements in this respect. Statements such as “more color schemes”, “few words are best”, or “less disorder” may be viewed as less exact, and thereby confuse more potential solvers, despite the fact that such statements contain more quantifiers.

As for Model 3, this study hypothesizes that the influence of linguistic style can be moderated by task reward in crowdsourcing task descriptions. Here the p-value was found to be less than 0.001, and the model was found to have explained 49.4% of the variance in the dependent measure ($R = 0.707$, $R^2 = 0.499$, $R^2_{adj} = 0.493$). According to these results, it can be concluded that the amount of reward offered has a significant regulatory effect on negative words, intellectual property, and first-person pronouns. However, the moderating effect that comes from the reward amount on negative words is found to be contrary to this study’s assumptions. One possible explanation for this is that when the reward offered for a task is high, and even though potential solvers will care less about the reliability of the employer as a result, the solvers will regardless hope to get more information from the task description in order to help them complete the task. In this, a greater number of negative words in the task description will represent clearer task requirements. As such, the reward amount will exhibit a positive effect in regulating the effect of negative words.

In processing this study’s sample data, it was difficult to accurately measure the frequency of articles due to fundamental differences between the Chinese and English languages. At the same time, the task descriptions in the study’s sample dataset were found to contain few anxiety words. It must be noted that these factors may have affected the outcome of the study’s final regression, obscuring the true impact of these variables on crowdsourcing participation. However, further investigation herein goes beyond the scope of this study. These considerations should be addressed in future research projects.

6. Discussion

6.1. Key Findings

In this study, H1(a), H3(a), H4, H5, H8(a) and H10 are accepted, however, for H2 and H6, the result is found to be contrary to this study’s assumptions. So, based on the research literature on sociolinguistics, URT, and source credibility theory, this study’s results and analyses show that solver participation will increase for crowdsourcing employers that utilize linguistic styles that are characterized by less quantifiers, less first-person pronouns, and more cognitive complexity, as well as for those employers who offer an intellectual property description in their task descriptions. In addition, the effect of first-person pronouns, words of cognitive complexity, and intellectual property descriptions can be moderated by the amount of the reward an employer offers their solvers. As such, the findings of this study stand in contrast to past research that has suggested the dynamics of crowdsourcing may be stable across variable task descriptions [32].

6.2. Theoretical Implications

This study contributes to the growing literature on online crowdsourcing contests. Firstly, past research on the topic is based primarily on switching cost theory, which suggested that a longer task description will increase learning costs and thereby reduce solver participation. Although some studies discuss the relationship between the level of detail found in such task descriptions as well as the credibility of the employer, these studies primarily utilize questionnaire surveys and lack empirical support from objective data. Moreover, other studies on the topic have not been refined to include factors related to linguistic style. In this study, however, it is argued that URT and source credibility theory are more suitable for elucidating and analyzing the detailed information hidden within crowdsourcing task descriptions. This research project has sought to explain how the linguistic style used by task employers affects crowdsourcing participation. Additionally, it is suggested that the effectiveness of a particular linguistic style depends on the amount of the reward employers will offer.

Secondly, this study contributes to the field in its application of research on linguistic style into a new area, namely crowdsourcing contests. Though many researchers have noticed the importance of linguistic style in many other areas, such as P2P lending [12], lie detecting in online dating [30], online medical advice expertise inferences [29], and so on, few of the findings

of these studies are relevant to the emerging and significant crowdsourcing context. Crowdsourcing is not a traditional medium of persuasion, but it also needs to attract solvers by reducing uncertainty and signaling greater employer credibility, so the effects of linguistic style are just as important to crowdsourcing as they are to other forms of online interaction.

Thirdly, this study notices the difference between users' initial attention and actual participation. On TaskCN crowdsourcing platform, potential solvers will first see a task's title, its reward, the task type, its duration, as well as some information on the task's employer. If the potential solver is interested in the task, he or she will click on the task details link to see the task description itself. According to the considerations they make after reading the task description, potential solvers will decide whether or not to take on the task. This study has shown that the linguistic style of the task description plays a significant role in transforming a prospective solver's attention into actual participation and commitment. Crowdsourcing task employers should take note of this in order to attracting more solvers for their future projects.

6.3. Practical Implications

From a pragmatic point of view, this study's research findings indicate that there are several strategies that employers can use to better describe their task requirements and attract greater solver participation. First, potential solvers will read the task description for the sake of reducing uncertainty. In this respect, employers should use wording that is characterized by more cognitive complexity and use less quantifiers. Second, offering high-quality information through their crowdsourcing task descriptions, employers can increase the perceived credibility of their tasks. To achieve this end, employers should use fewer personal pronouns and offer clear intellectual property declarations in their task descriptions. In general, the more detailed the task information provided by the employers is, and the more confident the potential solvers are in the employer, the more solvers will sign up and participate. Additionally, these effects can be moderated by the amount of the reward an employer offers to its solvers for completing a crowdsourcing task.

6.4. Limitations and Future Research

Despite its merits, this study leaves us with some unanswered questions. Firstly, a crowdsourcing task description is usually composed of three parts, the title, the text, and an attachment. Such attachments can contain useful linguistic information relevant to the field. Thus, another potential avenue for relevant future

research is in identifying and analyzing the additional instructions offered in crowdsourcing tasks. Secondly, a word count methodology was used in this study to measure the language style used by employers in their crowdsourcing task descriptions. Despite the accuracy and effectiveness of this methodology in processing large amounts of data, it regardless will miss the nuances of some more complex and underlying phenomena, so a multi-level approach for future research can be used to combine computerized content analysis with qualitative discourse analysis.

7. Acknowledgements

The authors are grateful for the financial support from National Key Research & Development Plan of China (2017YFB1400100) and the National Natural Science Foundation of China (71702206&71402136).

8. References

- [1] Abe, J. A. A., "Changes in Alan Greenspan's Language Use Across the Economic Cycle: A Text Analysis of His Testimonies and Speeches", *Journal of Language and Social Psychology* (30:2), 2011, pp. 212-223.
- [2] Berger, C. R., and Calabrese, R. J., "Some Explorations in Initial Interaction and Beyond: Toward a Developmental Theory of Interpersonal Communication", *Human Communication Research* (1:2), 1975, pp. 99-112.
- [3] Boudreau, K. J., and Lakhani, K. R., "How to Manage Outside Innovation", *Mit Sloan Management Review* (50:4), 2009, pp. 69-76.
- [4] Brabham, D. C., "Crowdsourcing as a Model for Problem Solving: An Introduction and Cases", *Convergence* (14:1), 2008, pp. 75-90.
- [5] Cheung, M. Y., Luo, C., Sia, C. L., and Chen, H., "Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations", *International Journal of Electronic Commerce* (13:4), 2009, pp. 9-38.
- [6] Davis, D., and Brock, T. C., "Use of First Person Pronouns as a Function of Increased Objective Self-Awareness and Performance Feedback", *Journal of Experimental Social Psychology* (11:4), 1975, pp. 381-388.
- [7] Geppert, J., and Lawrence, J. E., "Predicting Firm Reputation Through Content Analysis of Shareholders' Letter", *Corporate Reputation Review* (11:4), 2008, pp. 285-307.
- [8] Hancock, J. T., Curry, L. E., Goorha, S., and Woodworth, M., "On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication", *Discourse Processes* (45:1), 2007, pp. 1-23.

- [9] Howe, J., "The Rise of Crowdsourcing", *Wired Magazine* (14:6), 2006, pp. 1-4.
- [10] Johnson, W., "People in Quandaries; The Semantics of Personal Adjustment", *Science & Sanity: An Introduction to non-Aristotelian Systems and General Semantics*, Lancaster: Science Press 2nd edition Luthans F. (1977), *Organizational Behavior* 1946, 1946, pp. 144-146.
- [11] Knapp, M. L., Hart, R. P., and Dennis, H. S., "An Exploration of Deception as a Communication Construct", *Human Communication Research* (1:1), 1974, pp. 15-29.
- [12] Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., and Gorski, S., "Peer to Peer Lending: The Relationship Between Language Features, Trustworthiness, and Persuasion Success", *Journal of Applied Communication Research* (39:1), 2011, pp. 19-37.
- [13] Paivio, A., "The Role of Topic and Vehicle Imagery in Metaphor Comprehension", *Communication and Cognition* (19:3), 1986, pp. 367-387.
- [14] Paivio, A., "Dual Coding Theory: Retrospect and Current Status", *Canadian Journal of Psychology/Revue Canadienne De Psychologie* (45:3), 1991, pp. 255.
- [15] Pasupathi, M., "Telling and the Remembered Self: Linguistic Differences in Memories for Previously Disclosed and Previously Undisclosed Events", *Memory* (15:3), 2007, pp. 258-270.
- [16] Pennebaker, J. W., and King, L. A., "Linguistic Styles: Language Use as an Individual Difference", *Journal of Personality and Social Psychology* (77:6), 1999, pp. 1296.
- [17] Pennebaker, J. W., Mehl, M. R., and Niederhoffer, K. G., "Psychological Aspects of Natural Language Use: Our Words, Our Selves", *Annual Review of Psychology* (54:1), 2003, pp. 547-577.
- [18] Petty, R. E., Cacioppo, J. T., and Goldman, R., "Personal Involvement as a Determinant of Argument-Based Persuasion", *Journal of Personality & Social Psychology* (41:5), 1981, pp. 847-855.
- [19] Rude, S., Gortner, E., and Pennebaker, J., "Language Use of Depressed and Depression-Vulnerable College Students", *Cognition & Emotion* (18:8), 2004, pp. 1121-1133.
- [20] Simmons, R. A., Gordon, P. C., and Chambless, D. L., "Pronouns in Marital Interaction", *Psychological Science* (16:12), 2005, pp. 932-936.
- [21] Slatcher, R. B., Chung, C. K., Pennebaker, J. W., and Stone, L. D., "Winning Words: Individual Differences in Linguistic Style Among U.S. Presidential and Vice-Presidential Candidates", *Journal of Research in Personality* (41:1), 2007, pp. 63-75.
- [22] Soliman, W., and Tuunainen, V. K., "Understanding Continued Use of Crowdsourcing Systems: An Interpretive Study", *Journal of Theoretical and Applied Electronic Commerce Research* (10:1), 2015, pp. 1-18.
- [23] Stenmark, D., "Group Cohesiveness and Extrinsic Motivation in Virtual Groups: Lessons from an Action Case Study of Electronic Brainstorming", *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on, IEEE2002*, 2002, pp. 160-169.
- [24] Sun, Y., Wang, N., and Peng, Z., "Working for One Penny: Understanding Why People Would Like to Participate in Online Tasks with Low Payment", *Computers in Human Behavior* (27:2), 2011, pp. 1033-1041.
- [25] Sun, Y., Wang, N., Yin, C., and Zhang, J. X., "Understanding the Relationships Between Motivators and Effort in Crowdsourcing Marketplaces: A Nonlinear Analysis", *International Journal of Information Management* (35:3), 2015, pp. 267-276.
- [26] Tausczik, Y. R., and Pennebaker, J. W., "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods", *Journal of Language & Social Psychology* (29:1), 2010, pp. 24-54.
- [27] Ter Doest, L., Semin, G. R., and Sherman, S. J., "Linguistic Context and Social Perception: Does Stimulus Abstraction Moderate Processing Style?", *Journal of Language and Social Psychology* (21:3), 2002, pp. 195-229.
- [28] Terwiesch, C., and Xu, Y., "Innovation Contests, Open Innovation, and Multiagent Problem Solving", *Management Science* (54:9), 2008, pp. 1529-1543.
- [29] Toma, C. L., and D'Angelo, J. D., "Tell-Tale Words: Linguistic Cues Used to Infer the Expertise of Online Medical Advice", *Journal of Language & Social Psychology* (34:1), 2014, pp. 25-45.
- [30] Toma, C. L., and Hancock, J. T., "What Lies Beneath: The Linguistic Traces of Deception in Online Dating Profiles", *Journal of Communication* (62:1), 2012, pp. 78-97.
- [31] Vukovic, M., "Crowdsourcing for Enterprises", *Services-I, 2009 World Conference on, IEEE2009*, 2009, pp. 686-692.
- [32] Walter, T., and Back, A., "Towards Measuring Crowdsourcing Success: An Empirical Study on Effects of External Factors in Online Idea Contest", 2011.
- [33] Yang, Y., Chen, P. Y., and Pavlou, P. A., "Open Innovation: An Empirical Study of Online Contests", *International Conference on Information Systems, Icis 2009, Phoenix, Arizona, USA, December2009*, 2009, pp. 1