The Crowd on the Assembly Line: Designing Tasks for a Better Crowdsourcing Experience

Completed Research Paper

Sara Moussawi
Baruch College
CUNY Graduate Center
55 Lexington Avenue, New York
Sara.Moussawi@baruch.cuny.edu

Marios Koufaris
Baruch College
City University of New York
55 Lexington Avenue, New York
Marios.Koufaris@baruch.cuny.edu

Abstract

Leveraging crowd potentials through low paid crowdsourcing micro-tasks has attracted great attention in the last decade as it proves to be a powerful new paradigm to get large amounts of work done quickly. A main challenge for crowdsourcers has been to design tasks that trigger optimum outputs from the crowd while providing crowdsourcers with an experience that would attract them to the platform in the future. Drawing mainly from expectancy theory and the motivation through design of work model, we develop and test a theoretical framework to explore the impact of extrinsic reward valence and perceived task characteristics on perceived output measures in crowdsourcing contexts. We specifically focus on the impact of three crowdsourcing task dimensions: autonomy, skill use, and meaningfulness. Our findings provide support for our model and suggest ways to improve task design, use extrinsic rewards, and provide an enhanced crowdsourcing experience for participants.

Keywords: Crowdsourcing, Work redesign, Motivation Theories
Introduction

Crowdsourcing is the act of outsourcing tasks originally performed inside or outside the organization, to an indefinitely large, heterogeneous mass of potential workers through the internet, for the purpose of value creation, and for monetary or non-monetary incentives (Archak and Sundararajan 2009; Estellés-Arolas and González-Ladrón-de-Guevara 2012; Geiger et al. 2011; Hammon and Hippner 2012; Howe 2006; Kaufmann et al. 2011; Mason and Watts 2009; Sun et al. 2012; Zhao and Zhu 2012b; Zheng et al. 2011). Crowdsourcing can be viewed as a web-based, distributed, and no-contract form of outsourcing, i.e., the use of external agents to perform one or more organizational activities (Zhao and Zhu 2012a; Zhao and Zhu 2012b). It often involves a combination of different types of work or activities, such as crowd wisdom (collective intelligence), crowd creation, crowd voting and crowd funding (Geiger et al. 2011). Since random tasks tend not to be intrinsically enjoyable, crowdsourcing is relying more on financial remunerations, rather than intrinsic rewards (Mason and Watts 2009). Such remunerations vary from few cents, in the case of Amazon Mechanical Turk and similar low paid crowdsourcing platforms, to millions of dollars, like in the case of Kaggle Inc. (Ferguson 2013; Zhao and Zhu 2012a). The low-payments crowdsourcing model is more easily replicable than models based on intrinsic rewards. Consequently, it has attracted much attention as a substitute mode of production for traditional firms (Mason and Watts 2009).

Decades before the crowdsourcing phenomenon, in the mass production era, jobs were mainly characterized by repetitiveness, low skills requirements, limited social interaction, and predetermined use of tools and methods. Job providers knew that repetition leads to maximum proficiency levels at performing a task and thus results in very high levels of production. At the same time, however, mass production increases the sense of anonymity and decreases the sense of membership in a common work community (Walker and Guest 1952). This type of production helped companies achieve tremendous profits through economies of scale but at the cost of the workers’ satisfaction and performance. The crowdsourcing concept in paid contexts is based on the same idea of breaking down the task into its simple and basic elements, making it accessible to internet users and paying them small amounts of money (Hammon and Hippner 2012; Yang et al. 2008). While crowdsourcing was first perceived as the new collective way of innovation, open contribution, and creativity (Hammon and Hippner 2012), recent applications of crowdsourcing are shaping it into a web-enabled mass production platform. Based on this resemblance, this research study aims at better understanding the crowdsourcee’s perceptions of intrinsic as well as extrinsic motivators, and their association with perceived effort and performance in low-paid crowdsourcing contexts.

Prior research explored characteristics of the crowd in numerous contexts, sources of motivation, incentives for participation, and impact of varying monetary rewards in multiple settings on output level and quality (Brabham 2008; Ipeirotis 2010; Kaufmann et al. 2011; Mason and Watts 2009; Rogstadius et al. 2011; Sun et al. 2012; Zheng et al. 2011). For example, a recent study explored the relative impact of extrinsic versus intrinsic motivators on crowdsourcees and the time they spent on Amazon’s Mechanical Turk (Kaufmann et al. 2011), while other work examined the importance of intrinsic motivators on participation in crowdsourced contests (Zheng et al. 2011). In a separate study, the relationship between reward valence, perceived trust, perceived self-efficacy, and perceived effort was explored on Witkey, a Chinese marketplace (Sun et al. 2012).

In our study, we focus on crowdsourcing platforms for micro-tasks, such as Amazon’s Mechanical Turk, where crowdsourcees are invited to complete multiple short tasks that are themselves subtasks of larger jobs. We examine the crowdsourcees’ perceived task characteristics and their association with both perceived performance and perceived effort. Our study is unique in the current stream of research on crowdsourcing for three reasons. Though we don’t use Amazon’s Mechanical Turk, our results extend to all micro-task crowdsourcing platforms in general. Second, we explore perceived output measures in these contexts, and their association with perceived task attributes, which to our knowledge has not been examined before. Third, we try to
understand the crowdsourcees’ perceptions through exploring their motivations during the micro-task crowdsourcing experience.

More specifically, we explore the perceived value of extrinsic rewards and perceived task characteristics, and their relationship with perceived effort and perceived performance in micro-task crowdsourcing contexts where workers are required to complete simple and basic tasks for a low monetary reward per task. We develop and empirically test a theoretical model based on motivation theories, mainly the expectancy theory (Lawler III 1969) and the motivation through design of work model (Hackman and Oldham 1976). Our choices for the task attributes and theoretical basis are based on the assembly line paradigm we introduced earlier. The contribution of our research is of relevance to practitioners as well as researchers. First, a positive crowdsourcing experience will increase the likelihood of return to the platform to complete more tasks, most probably posted by the same crowdsourcer, i.e., the requester, the designer, or the organization that posts the tasks on the platform. Second, we explore specific task characteristics that can impact the crowdsourcing experience positively. Third, high perceived performance levels in intrinsically motivating contexts drive the crowdsourcees, i.e., the workers on crowdsourcing platforms, to work more, based on a cost-benefit analysis for the invested effort and the resulting perceived performance. Therefore, this study is not simply a re-evaluation of established theories in a new context, but it is also a contribution to theory through a novel approach towards a recent phenomenon. It provides new findings and insights about crowdsourcing that hold practical importance to various stakeholders and set the stage for interesting future research possibilities in the field.

In the next section we review the pertaining literature on motivation and crowdsourcing. We follow that with a presentation of our research model and hypotheses. We then describe our research methods: study design, construct measures, sample, and data analysis. Finally, we discuss the results and limitations of our research and provide directions for future research and implications for practice.

**Literature Review**

To be motivated to perform a task is to be moved, activated, and energized to do it (Ryan and Deci 2000a). Theories of management and work motivation differentiate between two types of motivation: intrinsic, where the outcome stems from the mere task and feelings related to its completion, and extrinsic, where a separate reward is associated with the task accomplishment. Deci’s Soma puzzle experiment in 1969, along with several experiments that followed, revealed that extrinsic monetary rewards cause loss of intrinsic interest in the activity (Deci 1975; Pink 2010). In addition, human beings have an inherent tendency to take interest in novel and challenging tasks, to exercise and extend their capacities, and to learn and explore (Deci 1972; Pink 2010; Ryan and Deci 2000a; Ryan and Deci 2000b). Sources of motivation for crowdsourcing activities vary by context. In crowdsourcing contests, intrinsic motivation, brought on by contests features, can be more important than extrinsic motivation for inducing participation (Zheng et al. 2011), while in crowdsourcing marketplaces, extrinsic motivation can impact the quantity of output and the time spent on the platform (Mason and Watts 2009; Rogstadius et al. 2011). In addition, the type or context of crowdsourcing projects may largely define the nature of the crowdsourcing community that participates in them. For instance the crowds in contest-based contexts such as Istockphoto, Threadless, TopCoder, InnoCentive, or Kaggle consist of more specialized communities than the crowds in marketplaces, such as Amazon Mechanical Turk or Witkey (Bonabeau 2009; Brabham 2008; Ipeirotis 2010; Kaufmann et al. 2011; Zheng et al. 2011).

In alignment with the economic approach to human behavior (Becker 1976), prior research has developed predictive models to estimate the worker reservation wage, which is the minimum wage a worker is willing to accept for the task. For instance, these models forecast the median reservation wage of a worker on Amazon Mechanical Turk to be around $1.38/hour (Horton and Chilton 2010). Furthermore, using a simple rule of thumb that states that each new prize should have around twice as high a marginal utility than the prize directly above it, crowdsourcing contests designers can easily determine the optimal prize structure (Archak and Sundararajan
However, in online labor markets, a trade-off between price and desired completion time based on formal pricing policies seems to be a better way than using simple rules of thumb. Finding the best ways to employ monetary rewards is necessary in paid crowdsourcing contexts as incorrect pricing can lead to task starvation and waste of used capital (Faridani et al. 2011).

In general, extrinsic rewards play a major role in recent crowdsourcing projects since crowdsourcing tasks are not always intrinsically rewarding (Mason and Watts 2009). However, extrinsic motivators can also have negative effects. First, in cases where the extrinsic reward is perceived to be the controlling factor, intrinsic motivation is likely to be reduced (Osterloh and Frey 2000). Second, if extrinsic rewards are given to participants for doing an intrinsically motivating work, and if the rewards are contingent on performance, their intrinsic motivation for the work will decrease (Deci 1975). Third, in crowdsourcing contexts, though higher payments impact the completion rate and time spent on the platform (Kaufmann et al. 2011; Mason and Watts 2009; Rogstadius et al. 2011), they do not lead to a more accurate performance (Rogstadius et al. 2011).

Despite the role extrinsic rewards play in crowdsourcing contexts, intrinsic motivators may be more successful for increasing participation and boosting performance. In fact, people will be most creative when they are predominantly motivated by the interest, enjoyment, satisfaction, and challenge of the work itself (Amabile 1988; Ryan and Deci 2000b; Shalley et al. 2004). A self-perpetuated cycle of positive work motivation is initiated by producing high-quality work of which the worker is proud (Hackman and Oldham 1976; Hackman and Oldham 1980). From that perspective, intrinsic motivation can be self-generating and self-propagating. In crowdsourcing markets, research has shown that intrinsic motivation triggered through a non-profit context cover story can improve the quality of output for the average task, sequences of task and tasks with varying levels of complexity (Rogstadius et al. 2011). Moreover, in the same context, reward valence and trust positively influence effort, while the relationship between self-efficacy and effort is moderated by task complexity (Sun et al. 2012). Furthermore, in crowdsourcing contests, intrinsic motivators such as contest autonomy, variety and analyzability are more important than extrinsic motivation for inducing participation (Zheng et al. 2011).

In this paper, we examine worker perceptions of extrinsic reward and task characteristics in micro-task crowdsourcing contexts. We develop and empirically test a theoretical model drawing on motivation theories, mainly expectancy theory (Lawler III 1969) and the motivation through design of work model (Hackman and Oldham 1976). Expectancy theory explains worker motivation based on a cognitive assessment of the task, effort and performance needed to reach the intended reward. The motivation through design of work model highlights job dimensions that help increase internal work motivation and quality work performance. We specifically explore three task dimensions: task autonomy, skill use and meaningfulness.

**Research Model**

An individual is considered to be intrinsically motivated to do a task when there is no apparent reward from the task. In that case, the completion of the task by itself and/or the feelings associated with the process constitute the reward (Deci 1975; Ryan and Deci 2000a; Ryan and Deci 2000b). By nature, human beings are curious and self-motivated, they want to learn and explore, and they want to develop their capabilities and master new skills (Pink 2010; Ryan and Deci 2000b). People who are intrinsically motivated are more excited and confident when compared to extrinsically motivated individuals, thus are more persistent and perform better (Ryan and Deci 2000b; Sheldon et al. 1997). We define intrinsic motivation as the inherent satisfaction from the task and extrinsic motivation as the separable outcome associated with the task accomplishment (Ryan and Deci 2000a). Furthermore, we consider extrinsic rewards as ones that satisfy one or more of the lower order needs in Maslow’s hierarchy of needs (Maslow 1943), while intrinsic ones fulfill higher order needs.

Given the low price per micro-task, switching costs on micro-task crowdsourcing platforms tend to be very low. Thus, what determines if the worker is going to return and do more tasks is mainly a positive crowdsourcing experience on the platform. In this case, the benefits or rewards brought
by the task completion should offset the cognitive demands of the task. The cognitive demands stand for the effort invested in the task at hand, while the rewards result from a need satisfaction, or a goal achievement. Therefore, on micro-task crowdsourcing platforms, workers are motivated by the rewards brought on by the experience as long as the effort-reward scale is slanted to the reward side. The expectancy theory of motivation can be applied to micro-task crowds contexts. This theory attributes a worker motivation to perform effectively to two factors: the effort-reward probability and the reward value. Effort-reward probability is the individual’s subjective probability that the amount of effort devoted effectively to a task predicts the amount of reward obtained from the task. This probability is further divided into two sub-probabilities: expectancy and instrumentality. Expectancy is the probability that effort leads to performance. Instrumentality is the probability that performance leads to reward. A reward satisfies one or more of the needs in Maslow's hierarchy: security, social, self-esteem and self-actualization needs. The reward value or valence refers to the individual’s perception of the reward to be obtained after completing the task (Lawler III 1969).

The reward from the task can be the money associated with it or an intrinsic reward satisfying certain needs for the crowdsourcer. Thus, monetary rewards used on Amazon Mechanical Turk are one possible compensation scheme. Since micro tasks on crowdsourcing platforms tend not to be intrinsically enjoyable (Mason and Watts 2009), financial rewards are playing an increasing role in these contexts. Previous work shows that, on these platforms, the perceived value of work is positively associated with the monetary reward amount (Mason and Watts 2009). We define extrinsic reward valence as the crowdsourcer's perceived value of the extrinsic reward(s) to be obtained after completing the task. We furthermore define perceived effort as the crowdsourcer’s self-assessment of effort intensity. While prior research extensively explored the actual impact of monetary rewards mainly confirming its impact on output quantity, results of a field survey on Witkey, a crowdsourcing marketplace in China, have shown a positive association between reward valence and perceived effort (Sun et al. 2012). That is, if the crowdsourcer highly values the extrinsic reward, this will most likely result in him/her choosing to make a significant effort in order to complete the task and get the reward. Therefore, while completing the task, the higher the perceived value of the extrinsic reward, the higher the perceived effort invested in the task since the worker will be spending longer time and working harder to reach the reward. The intensity of the association between the actual value of the extrinsic reward and its perceived value is mainly centered around the crowdsourcer’s needs and personal attributes.

H1: Extrinsic reward valence is positively associated with perceived effort.

Another way to impact the amount of invested effort is by increasing the intrinsic rewards of the task. Intrinsic rewards result from the fulfillment of one or more of the higher order needs in Maslow hierarchy. Based on the motivation through work design theory (Hackman and Oldham 1976), the presence of core task dimensions can lead to psychological states of experienced responsibility and meaningfulness, which consequently lead to the generation of positive cycles of intrinsic motivation, and increased performance on the job. Due to the high resemblance between micro-tasks crowdsourcing platforms and the traditional assembly line type of job, specifically through work characteristics such as repetitiveness, limited skill use, and predetermined use of tools and methods, we focus on three task characteristics that we believe are highly related to the micro-task crowdsourcing context: autonomy, skill use, and meaningfulness. In assembly line jobs, while repetition leads to high levels of expertise with the task at hand, low autonomy leads to lower job satisfaction. Even if the job is initially associated with a certain level of meaningfulness to the worker, over time, it loses its associated meaning and positive motivation. Diminishing job satisfaction and positive motivation eventually lead to a decline in worker performance levels. We aim to explore the workers’ experience on web-enabled mass production platforms through an evaluation of these three specific characteristics and their repercussions.

Task autonomy, i.e., the degree of freedom and independence while performing the job, determines the experienced responsibility for results (Hackman and Oldham 1976). Imposed goals diminish intrinsic motivation because they lead to diminished autonomy. When workers are given the authority to make decisions, they feel they are causal agents. Such feelings of personal causation, free choice and commitment are the essence of intrinsic motivation (Deci 1972; Deci
and Ryan 2000). We define perceived task autonomy as the degree to which the task provides the crowdsourcer with substantial freedom, independence and control over how to complete the task (Hackman and Oldham 1976; Hackman and Oldham 1980; Lawler III 1969). In crowdsourcing marketplaces, prior research has shown that task autonomy is more important to crowdsourcers than other intrinsic motivators (Kaufmann et al. 2011). This importance can result from a need to feel committed to the task or the goals behind it, or a need to feel in control of one’s actions, or even a need to feel responsible for the results and consequences. High perceptions of autonomy could therefore incentivize the worker to exert increased effort to finish the task at hand. As a result, it is likely that, in micro-task crowdsourcing platforms, a task high on autonomy will positively impact the perceived effort.

H2: Perceived task autonomy is positively associated with perceived effort.

The job characteristics model proposes that a job high on skill variety, i.e., the extent to which it requires the use of various skills, will lead to higher experienced meaningfulness of the work (Hackman and Oldham 1976). We define perceived use of skills as the crowdsourcee’s self-assessment of the extent to which the task requires him or her to use a variety of skills and abilities. The job must require the use of abilities and skills that the crowdsourcee values, in order to satisfy his or her sense of achievement and accomplishment, which can result in better performance (Lawler III 1969). In crowdsourcing contests, prior research has shown that when workers perform repetitive actions, where few skills are required, there is a negative impact on intrinsic motivation (Zheng et al. 2011). In crowdsourcing markets, research has shown that micro-tasks tend not to be intrinsically enjoyable (Mason and Watts 2009). Furthermore, in the same context, another study has shown that skill variety has the highest impact on the time workers spend on the platform when compared to other intrinsic motivators such as task autonomy and task identity. Additionally, workers view skill use as a very important motivator for their work (Kaufmann et al. 2011). Given that a job that includes different work activities is more enjoyable to the worker (Sims et al. 1976), it is more likely that the higher the perceived use of skills, the more intrinsically motivated and challenged the worker will be, which will impact the perceived effort levels invested in the task at hand.

H3: Perceived use of skills is positively associated with perceived effort.

Task meaningfulness refers to finding a sense of meaning in a task. Meaning is the value of a work goal, judged by the individual’s own principles and values (Thomas and Velthouse 1990). It is a fit between the work requirements and subjective beliefs (Hackman and Oldham 1980). We define perceived meaningfulness as the value brought by the crowdsourcing task itself. Organizational psychology researchers distinguish between intrinsic and extrinsic work value orientations. Intrinsic work value orientations are more leaning towards self-actualization and self-expression, i.e. achieving personal goals, making meaningful contributions to the societal entourage, and developing one’s skills and competencies. On the other hand, extrinsic orientations are related to security and material gain (Vansteenkiste et al. 2007). Task meaningfulness is thus associated with intrinsic work value orientations. Consequently, it is expected to satisfy higher order needs of self-expression and self-actualization on Maslow’s hierarchy of needs (Lawler III 1969). Prior research on Amazon Mechanical Turk has shown that for equal payments, meaningful contexts attract more workers than non-meaningful contexts (Chandler 2013). We believe that the increased perceived meaningfulness of the tasks will increase the crowdsourcee’s involvement. Therefore, s/he will be spending more time and devoting more effort to complete more tasks. Thus we expect task meaningfulness to positively impact the perceived effort in micro-task crowdsourcing contexts.

H4: Perceived task meaningfulness is positively associated with perceived effort.

The effort-performance probability, i.e. the expectation that more effort would lead to better performance, and the performance-reward probability, i.e., the expectation that better performance will lead to higher rewards (intrinsic or extrinsic), are the two main constituents of the expectancy theory of motivation (Lawler III 1969). Based on this theory, we argue that appropriate job design changes can impact intrinsic motivation positively because they may change the worker’s belief that rewards can be gained through better performance by putting
increased effort in the task at hand. We define perceived performance as the crowdsourcer's self-assessment of how well he or she performed on the task. We expect that when the crowdsourcer perceives his/her dedicated effort to be high, s/he will also perceive his/her performance to be higher.

H5: Perceived effort is positively associated with perceived performance.

One more possible way to increase crowdsourcer motivation is through the performance-reward probability. We propose that the perceptions of task characteristics and extrinsic rewards will also impact perceived performance independently from perceived effort. Based on the theory of motivation through the design of work, the mere presence of core job dimensions will lead to higher internal motivation for workers, which will cause them to perform effectively on their jobs (Hackman and Oldham 1976). First, increased task autonomy will result in increased responsibility for job outcomes (Hackman and Oldham 1976). Opportunities of self-direction and feelings of personal causation, free choice and commitment satisfy higher order needs in the Maslow hierarchy of needs and thus enhance intrinsic motivation (Deci 1972; Deci and Ryan 2000). These feelings and satisfied needs will furthermore create a cycle of self-perpetuating and self-rewarding intrinsic motivation leading to positive perceptions of performance levels. Since task autonomy is more important to crowdsourcers than other intrinsic motivators (Kaufmann et al. 2011), it is more likely that it will have a positive impact on perceived performance. That is, when the task at hand enhances feelings of commitment, independence, control and/or responsibility, it will consequently enhance the perceived quality of output.

H6: Perceived task autonomy is positively associated with perceived performance.

Second, according to the theory of motivation through job design, the use of various skills on the job leads to a psychological state of experienced meaningfulness on the work, which results in increased motivation and performance levels (Hackman and Oldham 1976). A job that requires workers to exercise their skills is more enjoyable to them (Sims et al. 1976), and contributes to higher levels of internal motivation through satisfying feelings of achievement and accomplishment (Lawler III 1969). The satisfaction of these higher order needs will help create cycles of positive motivation resulting in a positive perception of work outcomes. Prior research has shown that on Amazon Mechanical Turk, skill variety is more significant to workers than other intrinsic motivators (Kaufmann et al. 2011). Overall, we expect that the more workers feel that their skills are being used, the better they will expect the outcome of their work to be. In addition, we expect that a task designed to match the worker’s set of valued skills may increase his/her expectations of intrinsic rewards gained through his/her better performance. These rewards could be feelings of achievement, self-development, and/or accomplishment. Therefore, we propose that the stronger these perceptions of used skills are, the greater the perceived performance will be.

H7: Perceived use of skills is positively associated with perceived performance.

Third, achieving personal goals, improving one’s competencies and knowledge, and/or contributing to the society will create positive cycles of intrinsic motivation satisfying higher order needs in Maslow’s hierarchy. Meaningfulness can be a by-product of satisfying other needs (self-direction and self-accomplishment) or a direct result from the task and its completion (self-actualization and self-expression). Either way, it is associated with intrinsic work motivation rather than material gains and extrinsic motivation. These feelings create a self-actualizing, and self-rewarding atmosphere of positive intrinsic motivation leading workers to perform effectively on their jobs (Hackman and Oldham 1976), and perceive their work outcome in a positive manner. Therefore, irrespective of the effort invested, due to the positive attitudes created by these positive intrinsic motivation cycles, we expect that the higher the perceived task meaningfulness is, the higher the perceived performance in micro-task crowdsourcing contexts will be.

H8: Perceived task meaningfulness is positively associated with perceived performance.

Finally, the perceived value of the extrinsic reward varies as a result of the workers’ needs, goals and sources of motivation. While the average reward of $0.30 on Amazon Mechanical Turk might
not have a high value for some workers, the same reward will be greatly valued by others. Prior studies in Amazon Mechanical Turk detected an anchoring effect, where higher payments lead to higher perceived value of the work (Mason and Watts 2009). Based on the performance-reward probability of the expectancy theory of motivation (Lawler III 1969), we argue that a high extrinsic reward valence will increase the perceived performance during task completion. That is because if a worker perceives the extrinsic reward as having high valence, he/she will justify that by believing that their performance was also of high quality and deserving of the significant extrinsic reward.

H9: Extrinsic reward valence is positively associated with perceived performance.

Our model is presented in Figure 1 below.

**Research Methods**

**Study Design**

The crowd has a complex structure that varies by crowdsourcing genre (Brabham 2008; Ipeirotis 2010). Workers also vary by age, education, country, employment status (Ipeirotis 2010). As a result, a better understanding of task characteristics seems reliant on controlling some of these external factors to a certain extent. For this reason, we simulated a real crowdsourcing experience by creating our own crowdsourcing platform. We recruited student subjects, asked them to visit our crowdsourcing site, to complete crowdsourcing tasks, and then fill out a questionnaire.

For the purpose of this study we adapted a set of crowdsourcing micro-tasks from Amazon Mechanical Turk (Amazon 2012). Our subjects began by reading instructions that explained to them the nature of crowdsourcing tasks in general and our tasks in particular. We had two types of micro-tasks: classification and ownership identification. Classification micro-tasks required participants to determine what a company does and classify it into the appropriate industry category from a predetermined list. To complete this step, the subjects had to quickly review the information displayed in a screenshot of the home page of the business website. Ownership identification micro-tasks required determining whether the company is a subsidiary of another company. We first explained to subjects what a subsidiary is and how to read Google search
results based on a predetermined search query. They then determined if the company is a subsidiary by quickly reading a screenshot of the first page of results from a Google search that was provided to them. Tasks were designed to be quick and to allow for high variance in measures for autonomy, meaningfulness, and use of skills. The participants were presented with each micro-task in the same order. Since this is a cross-sectional study, maintaining the same order of tasks ruled out any possible order effects. After completing the micro-tasks, participants filled out a questionnaire. At the end, we collected descriptive data about them such as age, gender, education and current work, if any.

Before running our study, we conducted a pilot test with 50 participants. We asked workers for their comments and performed some preliminary tests to ensure reliability and validity of our measures. Consequently, we performed some minor changes to our tasks, and questionnaire items.

**Construct Measures**

All scales used in our study were adapted from prior relevant studies. In order to fit the research context, wording modifications were applied. All measures use a seven-point Likert scale. Task autonomy and perceived use of skills measures were adapted from Hackman and Oldham (1974). Items to measure task meaningfulness were adapted from Spreitzer (1995). Extrinsic reward valence and perceived performance measures were adapted from Sanchez et al. (2000). Perceived effort measures were adapted from Sun et al. (2012). All items in our measurement instrument are available in the Appendix.

**Sample**

Participants were undergraduate college students at a Northeastern university in the U.S. 450 subjects volunteered to participate in this study over a period of four weeks. The response rate was 94% resulting in 422 collected questionnaires. The exclusion of 18 invalid questionnaires resulted in a total of 404 complete and valid ones for data analysis. The participation was voluntary, and students received course credit upon completion of the tasks and questionnaire. About 50% of the subjects were 18 to 20 years old and 24% were 21 to 22 years old. Around 29% work full-time and 2% part-time.

**Data Analysis**

We conducted our data analysis in two steps. First, we examined the properties of the measurement scales. Second, we tested our model using partial least squares (PLS) structural equation modeling.

First, we performed a confirmatory factor analysis. We conducted a principal components analysis with oblimin rotation, since our constructs were expected to be correlated. Based on the results, presented in Table 2, all loadings were significant and greater than 0.70 indicating adequate convergent validity. To ensure discriminant validity, we used the square root of the average variance extracted (AVE) method (Fornell and Larcker 1981). Square roots of the average variance extracted are displayed in the diagonal cells of the correlation matrix in Table 1, and are greater than off-diagonal values. We tested construct reliability using Cronbach’s alpha. As a result of our reliability test, we dropped one reversed item from the perceived use of skills scale and three items from the task autonomy scale. Though having more items for the same construct results in greater confidence in our results, we strongly believe that the items we were left with for both constructs tap into the constructs we intend to measure and adequately cover the domains of interest. All final Cronbach’s alpha values were above 0.7, as shown in Table 1, indicating good reliability for all scales (Fornell and Larcker 1981). We also tested for common method bias using Harman’s single factor test (Harman 1976). Our principal components analysis showed that no single factor accounted for the majority of variance, thus indicating that common method bias was not significant in our data.
We used Partial Least Squares (PLS) with the SmartPLS software to test our model (Ringle 2005). Using PLS requires a minimum sample size of 40, i.e. ten times the largest number of formative indicators or structural paths directed to a dependent variable in our model (Chin 1998; Chin 2000). Thus our current sample size is appropriate for PLS analysis. The statistical significance of path coefficients was estimated based on the bootstrapping technique, as recommended by Chin (2010).

Results of the hypotheses testing are presented in Figure 2. H1 examined the positive effect of extrinsic reward valence on perceived effort. This hypothesis was supported ($\beta = 0.33$, $p= 0.001$). H2, H3, and H4 hypothesized the positive influence of task autonomy, skill use, and meaningfulness on perceived effort. All three hypotheses were supported: H2 ($\beta = 0.14$, $p= 0.001$); H3 ($\beta = 0.23$, $p= 0.001$); H4 ($\beta = 0.28$, $p= 0.001$). The path from perceived effort to perceived performance was statistically significant, thus H5 was supported ($\beta = 0.26$, $p= 0.001$). H6, H7, and H8 hypothesized the positive influence of task autonomy, skill use, and meaningfulness on perceived effort. Only H6 and H8 were supported: H6 ($\beta = 0.18$, $p= 0.001$); H8 ($\beta = 0.23$, $p= 0.001$). H7 showed a negative association between perceived use of skills and perceived performance, and thus was not supported ($\beta = -0.20$, $p= 0.001$). Finally, H9 examined the positive effect of extrinsic reward valence on perceived performance and was supported ($\beta = 0.26$, $p= 0.001$).

The variance explained by the antecedents of perceived performance was around 39%, while the antecedents of perceived effort explained around 33% of its variance.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Cronbach’s alpha</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrinsic reward valence (1)</td>
<td>0.82</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived task autonomy (2)</td>
<td>N/A</td>
<td>0.40</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived use of skills (3)</td>
<td>0.74</td>
<td>-0.11</td>
<td>-0.10</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived task meaningfulness (4)</td>
<td>0.84</td>
<td>0.05</td>
<td>-0.01</td>
<td>0.33</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived performance (5)</td>
<td>0.88</td>
<td>0.47</td>
<td>0.37</td>
<td>-0.10</td>
<td>0.27</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Perceived effort (6)</td>
<td>0.74</td>
<td>0.38</td>
<td>0.25</td>
<td>0.28</td>
<td>0.37</td>
<td>0.44</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2 Factor Loadings and Cross Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERV1</td>
</tr>
<tr>
<td>ERV2</td>
</tr>
<tr>
<td>ERV3</td>
</tr>
<tr>
<td>PTM1</td>
</tr>
<tr>
<td>PTM2</td>
</tr>
<tr>
<td>PTM3</td>
</tr>
<tr>
<td>PE1</td>
</tr>
<tr>
<td>PE2</td>
</tr>
<tr>
<td>PE3</td>
</tr>
<tr>
<td>PP1</td>
</tr>
<tr>
<td>PP2</td>
</tr>
<tr>
<td>PP3</td>
</tr>
<tr>
<td>PP4</td>
</tr>
<tr>
<td>PTA2</td>
</tr>
<tr>
<td>PUS1</td>
</tr>
<tr>
<td>PUS2</td>
</tr>
</tbody>
</table>
Discussion

This study explored the impact of extrinsic reward valence and three task-based intrinsic rewards (autonomy, skill use and meaningfulness) on perceived effort and perceived performance in micro-task crowdsourcing platforms. We developed and tested a theoretical model drawing mainly from expectancy theory and the motivation through design of work model. We simulated a micro-task crowdsourcing platform and collected data through a questionnaire that participants filled out after completing two micro-tasks. Our findings provide evidence that a task that is well designed from a motivational point of view will create a self-perpetuating cycle of positive motivation (Hackman and Oldham 1976) which will lead to higher perceived performance and perceived effort levels in crowdsourcing contexts. Our study makes several contributions to the IS literature.

As our first contribution, we propose ways for improved task design in crowdsourcing contexts. We show that designing tasks that motivate crowdsourcers to perform better is feasible. First, a feeling of free choice, independence, and self-direction while performing the task can lead the worker to exert increased effort. Prior research has shown that task autonomy is important for workers on crowdsourcing projects (Kaufmann et al. 2011). Our results confirm this finding by showing that increased autonomy will increase intrinsic motivation and thus will cause the worker to highly evaluate his/her performance. Perceived performance gives the worker assurance that through his/her hard work, s/he is able to reach the intended reward. Second, when the task is meaningful to the crowdsourcee, i.e., helps achieve a work goal, develop skills, increase knowledge, and/or contribute to the society, the crowdsourcee’s involvement in the task at hand is greater. This reflects a strong positive cycle of intrinsic motivation that results in the worker producing more and better output. Prior studies have shown that intrinsic motivation triggered through a non-profit context cover story can improve the quality of output (Rogstadius et al. 2011). Our study confirms and extends these prior findings by showing a positive association between perceived task meaningfulness and perceived performance and effort.

While we show that task autonomy and meaningfulness positively impact perceived effort and performance, our findings are not conclusive for the impact of perceived use of skills on perceived outcome measures. We found that skill use has a positive impact on perceived effort and a negative impact on perceived performance. One possible explanation could be that a worker who perceives the task as requiring more skill use will be devoting additional effort to complete the task. But the higher the perceived level of needed skills, the less confident s/he will be while completing the task. Low levels of confidence will affect the way the worker perceives her performance negatively. Thus, our findings aren’t decisive concerning the level of skill use in crowdsourcing tasks. Further research could investigate possible mediating factors for the relationship between skill use and outcome.

Our study also contributes to the literature by providing a different way to view extrinsic rewards in crowdsourcing contexts. Prior field studies on Amazon Mechanical Turk have shown that higher monetary incentives increase participation, as well as quantity but not quality of output (Mason and Watts 2009; Rogstadius et al. 2011). Our results show that it doesn’t matter how big the actual extrinsic reward is, what matters is how the worker perceives it. We showed that the more the worker values the extrinsic reward, the better s/he will evaluate his/her performance and the more effort s/he will invest in the task at hand. Prior studies detected an anchoring effect, where higher payments lead to higher perceived value of the work (Mason and Watts 2009). We propose a different route to increase the work value through task attributes, and a new standpoint from which to view and use monetary rewards. Therefore, the value of the extrinsic reward has a contextual association for the crowdsourcee, and designers need to take that into consideration when setting rewards for crowdsourcing tasks.
As a final contribution to the IS literature, our study indicates that better designed tasks can result in an enhanced crowdsourcing experience for the participants. We found that intrinsic motivation boosts perceived performance and effort. That is, a task high on autonomy and meaningfulness will create a cycle of positive work motivation through satisfying needs of self-actualization and self-expression and eliciting feelings of personal causation, free choice and commitment, which are at the core of intrinsic motivation (Deci 1972). Our results suggest that designing crowdsourcing tasks high on intrinsic motivation is promising for the crowdsourcer and the crowdsourcee. In low-paid micro-task crowdsourcing contexts, a low turnover rate could increase proficiency of workers at doing the tasks. Also, a positive cost-benefit assessment will increase his/her satisfaction with the platform, the task and the reward. The worker’s increased engagement will cause him/her to do more tasks and be willing to return to the platform to perform more work posted by the same crowdsourcer.

Our findings also have practical implications to providers of crowdsourcing platforms, crowdsourcers, and crowdsourcees. First, providers of crowdsourcing platforms should match worker skills to the micro-tasks they work on. On Amazon Mechanical Turk, micro-task creators have the option to customize worker requirements to fit their needs. For instance, a requester can control the proficiency of workers choosing their tasks by specifying the minimum threshold for the number of previously approved micro-tasks, or more easily by choosing to work with “Master” workers who demonstrate high accuracy in a specific type of micro-tasks (Amazon, 2012).

Second, crowdsourcers should clearly communicate the meaningfulness of the task to the workers. Meaningfulness in micro-task crowdsourcing contexts can be viewed as external or internal. External meaningfulness refers to the significance of the overall task or project to society or other people in general. If applicable, crowdsourcers can make clear that the overall task or project has extended benefits, such as helping research for a cure or archiving an intellectual treasure. Internal meaningfulness involves stressing the importance of the micro-task itself to the completion of the overall task or project. Crowdsourcers can make it clear that each micro-task completed by the worker is essential for the successful completion of the overall task or project.
Third, providers of crowdsourcing platforms and crowdsourcers should dig deeper to find the types and value of extrinsic rewards that matter to workers. Payments provided to workers on these platforms are relatively low. Therefore, finding out what matters to crowdsourcees is of crucial importance to improve their experience with the platform. Prior research studies varied financial reward in field experiment settings (Mason and Watts 2009; Rogstadius et al. 2011; Zheng et al. 2011). Crowdsourcers should draw on these findings to detect where completion rates start to drop off, and use these threshold figures to improve their compensation schemes.

Limitations

Despite several important contributions, our study has some limitations. First, our study was conducted in a simulated environment with undergraduate students instead of a real crowdsourcing platform with workers from different countries, with various ages, education levels, and occupations. We should note however, that the meaning participants give to the situation they are in along with their behavior play a bigger role in determining the generalizability of the study outcome than the sample representativeness (Berkowitz and Donnerstein 1982). The fact that our crowdsourcees didn’t visit the crowdsourcing site on their own is a limitation for our study. However, while it is possible that motivations and perceptions might differ when directed to the crowdsourcing site, we do not believe that the difference is large enough to compromise the generalizability of our results. Furthermore, students do represent a portion of the real crowdsourcing population. Thus, despite the fact that the real crowd often represents a much more complex environment, we resorted to a more controlled environment and a more homogeneous sample as it could help us better examine perceived task characteristics and outcome measures. Second, the actual reward was not a monetary reward as is the case of most crowdsourcing platforms, but rather course credit. Since course credit represents an extrinsic source of motivation, using it instead of financial incentives should not impact our results.

Directions for Research and Practice

Despite these limitations, this paper covers a gap in the literature by addressing four important issues: the crowdsourcing experience from the user perspective, the task design, the extrinsic reward valence and the components of positive intrinsic motivation. Our theoretical model draws on motivation theories to help understand crowdsourcees’ experience in micro-task platforms. We use expectancy theory to explain worker motivation based on a cognitive assessment of the task, effort, and performance needed to reach the intended reward. We refer to the motivation through job design theory to explore task dimensions that help increase internal work motivation and quality work performance.

One main assumption of the expectancy theory is that reward valence and effort-reward probability operate in a synergetic manner (Lawler III 1969). That is, if one exists without the other, the result will be that no motivation will exist. For the crowdsourcee, this is manifested in him/her wanting the reward associated with the task, and believing that the relationship between hard work and getting the reward exists. If s/he doesn’t want the reward, or doesn’t see the effort-performance, and performance-reward probabilities to be existent, s/he will not be motivated to do the task. Another main assumption we embrace in this study is the cost-benefit valuation that the worker starts with when s/he first chooses the task, and re-assesses all the way until the task is completed. Only when the result of this assessment is positive, i.e., perceived benefits are greater than perceived costs, would the worker do more tasks, and/or return to the platform. Our results confirm the cognitive valuation proposed by the expectancy theory that attributes a worker motivation to perform effectively to two factors: the effort-reward probability and the reward value. Furthermore our findings support prior field experiments results in Amazon Mechanical Turk that observed interaction effects between intrinsic and extrinsic motivations (Rogstadius et al. 2011).

Our model and results highlight the impact of task characteristics on perceived performance and effort. We proposed three task attributes: task meaningfulness, task autonomy and skill use. More
research is needed to explore other task features. The task identity, and the feedback provided by the task itself could provide interesting insight on improved task design in crowdsourcing contexts. Additionally, since our results on perceived use of skills were not conclusive, future research needs to explore possible mediators or moderators, such as task complexity and level of confidence, that might affect the relationship between skill use and output measures. Besides, other output measures need to be explored depending on the crowdsourcing platform genre. For instance, investigating possible design features in crowdsourcing communities that foster creativity and innovation could be very promising to help understand the motivational aspects on these platforms.

Our findings also have important practical implications. We started this paper by exploring an emerging resemblance between recent crowdsourcing platforms and the assembly line. Mass production work is characterized by repetitiveness, limited skill use, minimal social interaction and predetermined use of tools and methods. While repetition leads to optimum proficiency at performing the task, and while this type of production helps companies achieve remarkable profits, this comes at the cost of worker satisfaction. The new web-enabled mass production systems do not guarantee the same outcomes as the traditional assembly line for many reasons. First, the low switching costs prohibit these platforms from the gained proficiency acquired through repetitive task completion. Second, while the limited skill use helps increase proficiency, we find that it limits the engagement of the workers on crowdsourcing platforms. Third, we find that the more control the crowdsourcers have over their work, the more involved they will be. Consequently, web-enabled mass production systems may not be a sustainable option. Therefore, practitioners need to give utmost priority for designing better tasks that enhance output but not at the cost of workers’ satisfaction. Based on our results, this is a feasible scenario and a winning solution for both the crowdsourcer and the crowdsourcer.

**Conclusion**

Can better task design enhance the crowdsourcing experience for the crowdsourcer? Based on our findings, the answer to this question is yes. While our findings on perceived use of skills are not conclusive, we show that a task high on autonomy and meaningfulness will improve crowdsourcer’s engagement and perceived performance. Furthermore, we find that the perceived value of the extrinsic reward impacts both performance and effort positively. Our research study fills a gap in the literature by addressing several important issues: task design, reward valence, intrinsic motivation and their impact on perceived performance and effort. Based on our theoretical model and empirical results, we were able to propose possible ways to enhance task design and improve the crowdsourcing experience in low-payment micro-task crowdsourcing platforms.
References


Appendix: Measures

Extrinsic Reward Valence (Adapted from Sanchez et al. (2000))
ERV1: I would like to earn the course credit for completing the tasks.
ERV2: It would be good to earn the course credit for completing the tasks.
ERV3: I want to get the course credit.

Perceived Task Autonomy (Adapted from Hackman and Oldham (1974))
PTA1: I was able to do these tasks independently.

Perceived Use of Skills (Adapted from Hackman and Oldham (1974))
PUS1: The tasks required me to do many different things using a variety of skills.
PUS2: The tasks required me to use a number of complex or high-level skills.

Perceived Task Meaningfulness (Adapted from Spreitzer (1995))
PTM1: Working on such tasks is very important to me.
PTM2: Classifying organizations and exploring their ownership status is personally meaningful to me.
PTM3: Working on such tasks is meaningful to me.

Perceived Effort (Adapted from Sun et al. (2012))
PE1: I worked at my full capacity to finish the tasks.
PE2: I tried hard to finish the tasks.
PE3: I put a lot of effort into completing the tasks.

Perceived Performance (Adapted from Sanchez et al. (2000))
PP1: I expect that I did well enough on the tasks I just completed.
PP2: I believe I did well on the tasks I just completed.
PP3: I believe that I answered correctly a good number of tasks related questions.
PP4: I believe that I answered correctly more than half of the tasks related questions.