Computers in Human Behavior 54 (2016) 531–546

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Work experiences on MTurk: Job satisfaction, turnover, and information sharing

Alice M. Brawley*, Cynthia L.S. Pury

Clemson University, 418 Brackett Hall, 321 Calhoun Drive, Clemson, SC 29634, USA

ARTICLE INFO

Article history:
Received 27 May 2015
Received in revised form
20 August 2015
Accepted 24 August 2015
Available online 11 September 2015

Keywords:
Crowdsourcing
Technology in work
Changing nature of work
Job satisfaction
Amazon's Mechanical Turk
MTurk

ABSTRACT

Amazon's Mechanical Turk (MTurk) is an online marketplace for work, where Requesters post Human Intelligence Tasks (HITs) for Workers to complete for varying compensation. Past research has focused on the quality and generalizability of social and behavioral science research conducted using MTurk as a source of research participants. However, MTurk and other crowdsourcing platforms also exemplify trends toward extremely short-term contract work. We apply principles of industrial–organizational (I–O) psychology to investigate MTurk Worker job satisfaction, information sharing, and turnover. We also report the top best and worst Requester behaviors (e.g., building a relationship, unfair pay) that affect Worker satisfaction. Worker satisfaction was consistently negatively related to turnover as expected, indicating that this traditional variable operates similarly in the MTurk work context. However, few of the traditional predictors of job satisfaction were significant, signifying that new operational definitions or entirely new variables may be needed in order to adequately understand the experiences of crowdsourced workers. Coworker friendships consistently predicted information sharing among Workers. The findings of this study are useful for understanding the experiences of crowdsourced workers from the perspective of I–O psychology, as well as for researchers using MTurk as a recruitment tool.

Amazon's Mechanical Turk (MTurk) and other crowdsourcing platforms are harbingers of future extremely short-term contract employer–employee relationships. Crowdsourcing sites like MTurk represent changes to the organization of work at the organizational level and perhaps also the personal level (cf. Sauter et al., 2002) and exhibit trends in increased use of contingent workforces and short-term contracts (Cartwright, 2003). Crowdsourced work introduces novel relationships between employers and employees, effects of technology, facilitation of globalization (Ryan & Wessel, 2015), and ways of designing, producing, and conceptualizing work (Kittur et al., 2013). MTurk has a growing presence in the psychological literature as a source of research participants, and researchers in some fields, such as human computation, have examined work experiences on MTurk. We apply knowledge from long-term, in-person work relationships traditionally studied in industrial–organizational (I–O) psychology to the very short-term online work experiences of crowdsourcing. In the present study, we test antecedents of job satisfaction, information sharing, and turnover among MTurk Workers and report the top best and worst Requester practices (i.e., those that result in the highest and lowest average job satisfaction) described by MTurk Workers.

MTurk (mturk.com) is a crowdsourcing platform that enables Requesters to post Human Intelligence Tasks (HITs). HITs typically encompass tasks which require little to no training but which can only be performed by humans, not by computers, such as complex image categorization, psychological survey completion, and so on. Each HIT includes the Requester's name, the compensation offered by the Requester, and a short description of the work involved in the HIT. Workers can choose from the available HITs those that they want to complete. Once a Worker submits a completed HIT, the Requester may review the submission before approving it and paying the Worker. If the Requester deems a submission unsatisfactory (e.g., tagging photos incorrectly, apparent random responding to survey questions), then the Requester can reject a Worker's submission and deny payment. Workers build an approval rating on MTurk by successfully completing HITs. Workers' approval ratings are harmed when work is rejected, and they may be prevented from completing future HITs requiring some minimum approval rating.

* Corresponding author.
E-mail addresses: ABRAWLE@clemson.edu (A.M. Brawley), CPURY@clemson.edu (C.L.S. Pury).

http://dx.doi.org/10.1016/j.chb.2015.08.031
0747-5632/

Requesters also build a reputation, though not through official MTurk channels. There are a number of forums and sites where Workers share information about Requesters with other Workers (see e.g., Bederson & Quinn, 2011; Marshall & Shipman, 2013; Martin, Hanrahan, O'Neill, & Gupta, 2014; and Schmidt, 2015; for discussion of where Workers share information and what types of information are shared). Workers therefore may have information that helps determine which Requesters to – and not to – work for (Chandler, Mueller, & Paolacci, 2014). This system may provide some degree of regulation of Requester behavior and decision-making. However, while both Requesters and Workers cannot change their alphanumeric ID, Requesters on MTurk can change their displayed name on the platform (including in HIT listings) at any time. We also note that due to payment policies as well as apparent changes in MTurk procedures limiting approval of new international Workers, the majority of MTurk Workers are residents of the United States (US) and India. Current policy also requires that Requesters have a valid US address, which may limit the number of new non-US Requesters on the site.

Despite MTurk's increasing popularity for research in the field of psychology and its study as a work experience by researchers in fields such as human computation, study by the field of I–O psychology – which applies psychology to work experiences – has focused on MTurk as a source of research participants. For example, Barger, Behrend, Sharek, and Sinar (2011) recommended determining fair pay according to pay for similar HITs and depending on the time requirement of one's own HIT. Barger et al. also recommended the use of standard ethical research procedures (i.e., providing informed consent and maintaining confidentiality) on MTurk. In a recent focal article in *Industrial and Organizational Psychology*: *Perspectives on Science and Practice*, Landers and Behrend (2015) discussed the implications of MTurk as a convenience sample. Schmidt (2015) described the work experience of MTurk Workers from a participant observer perspective, but focused on the context of MTurk as a data source, rather than a work experience.

Researchers in I–O and other fields have compared MTurk samples to university and community samples (Behrend, Sharek, Meade, & Wiebe, 2011; Buhrmester, Kwang, & Gosling, 2011; Goodman, Cryder, & Cheema, 2012; Paolacci & Chandler, 2014) and examined other important issues for research conducted using MTurk, including measurement invariance and social desirability (Behrend et al., 2011), Workers' motivations for completing MTurk tasks (Buhrmester et al., 2011), nonindependence between MTurk samples (Chandler et al., 2014), HIT design and incentives (Chandler, Paolacci, & Mueller, 2013), and the comparability or viability of using MTurk for various tests and methods (e.g., Casler, Bickel, & Hackett, 2013; Holden, Dennie, & Hicks, 2013; Marshall & Shipman, 2013; Summerville & Chartier, 2012).

I–O psychology – a field that studies the psychology of work experiences – has not yet explicitly been applied to understanding the work experience of Workers on MTurk or other crowdsourcing channels. The earliest studies in the field of I–O psychology focused on large organizations such as the U.S. Army and Western Electric Company (Levy, 2006), and I–O has remained mostly focused on large, hierarchical organizations, with rigidly defined jobs (Tetrick, Slack, Da Silva, & Sinclair, 2000). However, organizations are becoming leaner and less rigidly defined (Cascio & Aguinis, 2011), and the field must be flexible in order to remain relevant as the nature of work and the workplace changes (Landy & Conte, 2009). As a first step in studying the work experience of crowdsourcing on MTurk, we will examine whether I–O findings hold true in this new context. Current psychological findings regarding work attitudes and behaviors may generalize to all types of work, not just long-term, single-employer employment relationships. In Section 1.1, we compare MTurk to traditional work and review past research examining crowdsourcing on MTurk as a work experience.

Traditional organizational research has found that job satisfaction may be predicted by a combination of individual differences and situational characteristics (e.g., Judge & Kammeyer-Mueller, 2012), while job satisfaction in turn predicts turnover (Brown & Peterson, 1993; Griffeth, Hom, & Gaertner, 2000; Tett & Meyer, 1993). We discuss the application of research on job satisfaction and turnover to the MTurk context in Sections 1.2 and 1.3, respectively. After adapting the construct definitions of job satisfaction and turnover for crowdsourcing on MTurk, we examined whether previously supported individual difference and situational antecedents predict job satisfaction and turnover among MTurk Workers. We also explore information sharing behaviors of MTurk Workers (discussed in Section 1.4), and whether the same antecedents that predict job satisfaction predict information sharing behaviors. Understanding the antecedents of job satisfaction, turnover, and information sharing among MTurk Workers will inform crowdsourcing researchers about this new type of work experience.

1. Previous research on MTurk work experiences

The Requester–Worker relationships on MTurk differ in many respects from traditional work arrangements. For example, Workers on MTurk can decide more easily and quickly who they will begin to and discontinue working for than in traditional employment. However, Requester–Worker relationships share some similarities with alternative forms of work, including telework, contracting, and temporary or contingent work. As with telework, Workers on MTurk complete tasks remotely. However, a likely difference between telework and MTurk is that Workers on MTurk could potentially complete tasks with no direct communication with a given Requester. Furthermore, MTurk Workers select among available tasks to complete based on their own preferences, rather than having tasks assigned to them or particular tasks for which they are regularly responsible, which may more likely occur with telework.

Working on MTurk is similar in some respects to contract work. In both instances, terms of payment are established prior to beginning any work, and payment is made when the work is completed to the employer's request. However, tasks on MTurk are likely more limited in scope than those completed by contracted employees. Last, working on MTurk shares the transient nature of temporary or contingent working arrangements. As with contracting, tasks on MTurk are likely more limited in scope than those obtained through typical temporary or contingent work arrangements, and may even be extremely short term – sometimes merely a few seconds – compared to typical temporary work.

Researchers in some fields and MTurk Workers themselves have recognized MTurk as a work experience. Many Workers list MTurk as their employment on Facebook (Gupta, Crabtree, Rodden, Martin, & O'Neill, 2014) and researchers have noted that Requesters are, if only briefly, "their subjects' employers" (Antin & Shaw, 2012, p. 4). Preliminary research in information science suggests that the I–O theory of person-job fit can be adapted to crowdsourcing markets (Schulze, Krug, & Schader, 2012), and computer science researchers found that performance-based pay mattered for tasks where time spent affected the quality of work (Ho, Slivkins, Suri, & Vaughan, 2015), similar to the Expectancy component of Vroom's theory of motivation (1964).

Many researchers have voiced concern about "occupational hazards" of working on MTurk (Silberman, Irani, & Ross, 2010). For example, Worker invisibility promotes companies that use MTurk (Irani, 2015a, 2015b; Irani & Silberman, 2013) while devaluating

Workers (Martin et al., 2014). Workers themselves may face identity challenges in the absence of a traditional organizational structure (Lehdonvirta & Mezier, 2013), presence of multiple Requesters, and relatively small scope of tasks completed (Zittrain, 2008). While Workers may perceive MTurk Requesters as fair or fairer than traditional employers (Horton, 2011), Workers typically have limited information about the ways their work will be used by Requesters (Zittrain).

Furthermore, Workers may increasingly rely on MTurk as a full-time job or as a source of income used to meet basic needs (Ross, Irani, Silberman, Zaldivar, & Tomlinson, 2010), and Worker wages have been discussed in a number of fields, including I–O psychology. Researchers have recommended that Requesters price tasks based on time; design tasks so as to not waste Worker time; disclose and follow terms of payment; and have a grievance process (Bederson & Quinn, 2011). I–O psychology researchers have recommended determining fair pay according to pay for similar HITs and depending on the time requirement of one's own HIT, with a recommendation for pay ranging from near minimum wage to 50 cents for an hour of work (Barger et al., 2011). These recommendations may align with the suggested considerations that MTurk Workers require less coordination and travel than lab-based research participants, have more control over their entry and exit in a research study, and may be less likely to depend on MTurk for primary income (Mason & Suri, 2012), though this conclusion about Workers may be debated (cf. Ross et al.).

In response to the hazards of working on MTurk, external platforms such as Turkopticon (Irani & Silberman, 2013) – for Workers to review Requesters – and Dynamo (Salehi et al., 2015) – for Workers to gather support for Worker initiatives – have been developed to promote Worker visibility. Other researchers have developed a framework for sustainable crowd work that considers both Worker needs (e.g., feedback) and Requester needs (e.g., quality control), and thereby addresses the future state of crowd-sourced work (Kittur et al., 2013).

The work experiences of MTurk Workers are diverse. Workers' motivations for working on MTurk vary considerably (Antin & Shaw, 2012; Buhrmester et al., 2011). There may be cross-cultural differences among Workers: Workers from India may feel more satisfaction and pride about working on MTurk than do Workers from the United States (Gupta et al., 2014). There may also be tenure-related differences among Workers: more experienced Workers may view newer Workers who will work for displeasing Requesters similar to the way unionized workers view those who work during strikes (Staffelbach et al., 2014). These differences, perhaps among others not yet identified, indicate that – like traditional employees – MTurk Workers are not a uniform population with uniform work experiences.

1.1. Why study work experiences on MTurk?

Crowdsourcing – and MTurk in particular – has grown in popularity among both Workers and Requesters. According to mturk-tracker.com (Ipeirotis, 2010a), MTurk revenues have grown consistently since 2009. MTurk is a source of inexpensive human computation (Buhrmester et al., 2011), making it a valuable outlet for Requesters. It is also an accessible source of work for individuals of various cognitive abilities, physical abilities, and locations (Paritosh, Ipeirotis, Cooper, & Suri, 2011; Schriner & Oerther, 2014; Zyskowski, Morris, Bigham, Gray, & Kane, 2015), making MTurk a valuable opportunity for many Workers, as well. MTurk has also grown in popularity in many fields of research as a method for sampling participants. To begin to study the work experiences of MTurk Workers using principles of I–O psychology, we begin with a broad study of job satisfaction based on results from traditional work settings: We examine dispositional, situational, and interactive factors as predictors of job satisfaction and information sharing, and we examine job satisfaction as a predictor of turnover.

1.2. Job satisfaction

Traditional I–O psychology research has found that job satisfaction may be predicted by both individual differences and situational characteristics (e.g., Judge & Kammeyer-Mueller, 2012), resulting in three broad categories of predictors of job satisfaction: dispositional, situational, and dispositional-situational (or person-situation) interactive factors. Predictors of each type have been supported in the research literature on job satisfaction in traditional employment, but each of these approaches may manifest and operate differently in the MTurk work context.

1.2.1. Dispositional approach

The dispositional approach to job satisfaction hypothesizes that stable individual differences – such as personality traits – predict job satisfaction, as job satisfaction may be relatively stable over time and across employers (e.g., Staw & Ross, 1985). Several dispositional factors have been supported as predictors of job satisfaction, including trait positive and negative affectivity (Cropanzano, James, & Konovsky, 1993; Judge, Heller, & Klinger, 2008; Shaw, Duffy, Jenkins, & Gupta, 1999), core self-evaluations (CSE; Judge et al., 2008; Lemelle & Scielzo, 2012; Brown & Peterson, 1993), as well as the Big Five personality traits (Judge et al., 2008; Templer, 2012).

However, the effectiveness of individual differences in predicting job satisfaction may be moderated by the work context. For example, diligence predicts job satisfaction among traditional employees, but not among teleworkers (O'Neill, Hambley, Greidanus, MacDonnell, & Kline, 2009). Certain traits may not be required or satisfied (e.g., diligence when unmonitored, extraversion when alone) by work in unmonitored and perhaps solitary work settings such as MTurk. It is possible that MTurk Workers exhibit systematic differences in personality traits such as extraversion and self-esteem when compared to student and community samples (Goodman et al., 2012). Such differences may alter the utility of individual differences for predicting job satisfaction among MTurk Workers.

Specifically, we studied three antecedents of job satisfaction as dispositional factors in the MTurk work context, as they are relatively stable within an individual across many HITs or Requesters. First, the Big Five personality factors and CSE will be studied as dispositional variables. Both the Big Five and CSE have been shown to predict job satisfaction in traditional work settings (e.g., Judge et al., 2008), and these individual differences may also predict job satisfaction among MTurk Workers. One additional relatively stable individual difference was investigated: coworker friendships. As mentioned previously, information sharing among Workers is an important tool for Workers to determine which Requesters to work for; this network of Workers may also provide social support or other positive benefits to Workers (e.g., Salehi et al., 2015; Schmidt, 2015), such as increased job satisfaction.

1.2.2. Situational approach

The situational approach to job satisfaction hypothesizes that situational features predict job satisfaction. That is, features of the work and work environment may best predict job satisfaction. In particular, job characteristics, the social environment, leadership, and organizational practices (see Judge & Kammeyer-Mueller, 2012) have been identified as common predictors of job satisfaction in traditional work settings.

Situational features may take on a different form or meaning in a new work context like MTurk. For example, Requester responsiveness and interactions may serve as contextual factors (Martin et al., 2014) in the absence of formal leadership or management systems. Job characteristics on MTurk may also differ systematically from those of traditional work; in particular, remote work may be more autonomous and less complex (O'Neill et al., 2009).

Specifically, we studied two antecedents of job satisfaction as situational factors in the MTurk work setting, as they require input mainly from the Requester or result mainly from the relatively objective features of the HIT. First, job characteristics (Hackman & Oldham, 1974) were studied as situational variables. An example job characteristic is dealing with others, or the degree to which the work requires interacting with other people. Job characteristics have been shown to predict job satisfaction in traditional work settings (Loher, Noe, Moeller, & Fitzgerald, 1985; Williams, McDaniel, & Nguyen, 2006) and may also predict job satisfaction in the MTurk work context. We also studied perceived organizational support – operationalized as perceived Requester support – as a situational antecedent of job satisfaction among MTurk Workers.

1.2.3. Interactive approach

Last, the interactive approach integrates both dispositional and situational factors as predictors of job satisfaction, with the effect of each predictor dependent on the other (Judge & Kammeyer-Mueller, 2012). For example, individuals can perceive the same workplace situational characteristics differently depending of their dispositions. Past research has supported a variety of such interactive (i.e., person-situation) factors – such as perceptions of role ambiguity and role conflict (Brown & Peterson, 1993) – as predictors of job satisfaction.

Such interactive factors may also differ in form or function on MTurk. MTurk Workers, as mentioned earlier, may possess a unique set of dispositional factors (Goodman et al., 2012), and MTurk is a novel context for work. The distinctiveness of both factors may result in novel interactive effects that predict job satisfaction.

Specifically, we studied four antecedents of job satisfaction as interactive factors in the MTurk work setting, requiring input or appraisal from both the Worker (i.e., person) and the HIT or Requester (i.e., situation). First, positive and negative affective states will be studied as interactive variables. Trait positive and negative affectivity have been linked to job satisfaction in traditional work arrangements (Cropanzano et al., 1993; Judge et al., 2008; Shaw et al., 1999). However, we studied affective reactions to specific Requesters or HITs, thus incorporating both person – Workers' appraisal – and situation – features of the Requester or HIT – factors into this antecedent of job satisfaction. Similarly, we studied psychological (i.e., implicit) contract fulfillment, intrinsic motivation, and pay satisfaction as reactions to specific Requesters or HITs and therefore as interactive variables that predict job satisfaction.

1.2.4. Job satisfaction hypotheses

The present study uses the critical incidents technique to examine job satisfaction among MTurk Workers. We asked Workers to describe satisfying or dissatisfying (i.e., highly salient positive or highly salient negative) experiences with Requesters and HITs on MTurk. We then examined each Worker's job satisfaction in the positive or negative experience. Based on the bimodal nature of this technique (i.e., critical positive or negative events), we propose that different antecedents will predict job satisfaction in positive experiences than will predict job satisfaction in negative experiences.

Negativity bias (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001) holds that negative information is more salient than positive information. Therefore, in negative work experiences on MTurk, compared to positive experiences, Workers should perceive comparatively more salient situational information. When there is more salient situational information – i.e., in negative events, as suggested by negativity bias – situational strength theory suggests that situational factors will be stronger predictors of behavior (Cooper & Withey, 2009; Meyer & Dalal, 2009; Snyder & Ickes, 1985). Expanding this prediction to an attitude, we propose that job satisfaction in negative experiences will be most strongly predicted by situational factors.

Hypothesis 1a. Job satisfaction in negative experiences will be more strongly predicted by situational factors than by dispositional or person-situation interactional factors.

Conversely, because there is less salient situational information in positive events, dispositional factors are predicted to be the strongest predictors of job satisfaction in positive events.

Hypothesis 1b. Job satisfaction in positive experiences will be more strongly predicted by dispositional factors than by situational or person-situation interactional factors.

1.3. Turnover

In traditional I–O psychology research, job satisfaction is negatively related to turnover (Brown & Peterson, 1993; Griffeth et al., 2000; Tett & Meyer, 1993). Job satisfaction may also be a mediator between other variables – such as dispositional affectivity – and turnover (Cropanzano et al., 1993). On MTurk, Workers have fewer barriers to turnover than do traditional employees. To quit a traditional job, employers may require a letter of resignation, the return of work-issued items, and the completion of an exit interview, while the employee must ensure that they have a new job offer or savings to cover costs of living while searching for a new job. A number of reasons and processes might accompany or lead to the decision to quit a traditional job (Maertz & Campion, 2004). In contrast, MTurk Workers can simply stop working on a HIT and avoid accepting new HITs from a particular Requester at any time they choose, with assurance that other Requesters will post HITs that the Worker can complete. Since there is little to no barrier to the decision to stop working for a particular Requester, an MTurk Worker's job satisfaction with a particular Requester may strongly predict Requester-specific turnover by that Worker. Based on research in traditional work settings and supported by the reduced barriers for MTurk Workers to quit immediately when desired, we expect that job satisfaction will be negatively related to turnover in the context of work on MTurk.

Hypothesis 2. Specific job satisfaction (i.e., with a particular Requester or HIT described by the Worker) will be negatively related to turnover specific to that Requester.

1.4. Sharing information

MTurk Workers can choose to work for Requesters based on their reputation among Workers. This information is available through a number of external forums and sites where Workers review Requesters and HITs. In the absence of formal regulations of Requester behavior, these resources provide Workers with information to decide which Requesters to (and not to) work for (Chandler et al., 2014). There may be strong norms among Workers – especially more experienced Workers – to refuse work to unfair

Requesters (Staffelbach et al., 2014). Because Workers may rely on shared information to decide which Requesters to work for, it may be important for Requesters to maintain a positive reputation among Workers. Unfair treatment to a small number of Workers can be quickly disseminated to many MTurk Workers, leaving a Requester little to no workforce.

To understand how to improve and maintain positive Requester reputations among Workers, we examined three research questions using both quantitative and open-ended data. First, is the pattern of information sharing predictable from the antecedents discussed previously? If so, emphasizing such factors in HIT design and implementation may positively influence the information shared about oneself as a Requester.

Research Question 1: Do any of the antecedents of job satisfaction – dispositional, situational, or interactive factors – predict information sharing by MTurk Workers?

To more thoroughly understand positive information sharing and reputation building, we further examined whether the predictors of information sharing differed across positive and negative work experiences.

Research Question 2: If any of the study variables predict information sharing, do different variables predict information sharing about positive or negative events?

The final research question examined Workers' open-ended descriptions of positive and negative work experiences, and how these experiences affect job satisfaction.

Research Question 3: What Requester practices are associated with satisfying and dissatisfying work experiences?

2. Method

2.1. Procedure

Workers were recruited via two separate HITs posted on MTurk, with one HIT each posted for residents of the US and residents of India. After viewing an informed consent page and consenting to participate, Workers were randomly assigned to one of two conditions that were identical except for a single narrative question that asked Workers to describe a time when he/she was either *very satisfied* (i.e., Positive Event condition) or *very dissatisfied* (i.e., Negative Event condition) with the way a Requester treated him/her or the HIT he/she completed for a Requester. This event was the focus of several additional questions.

2.2. Participants

We received 255 US responses and 258 Indian responses. All Workers were compensated $3.75 for the HIT. We determined this rate by choosing to pay approximately 1.5 times the US federal minimum wage ($7.25/hour) for the amount of time the HIT was estimated to require (15–20 min). We eliminated data that included random responding ($n = 81$; see attention-checking item description in Section 2.3.3) and responses that did not describe a particularly good or bad experience ($n = 75$), resulting in a final sample of 225 US Workers and 132 Indian Workers. Complete demographics of the samples are reported in Table 1.

2.3. Measures

All measures in this study were modified to suit MTurk: references to organizations, coworkers, and jobs were revised to refer to Requesters, other Workers, and HITs, respectively. Sample items are provided, and complete measures are available upon request. Correlations and reliabilities for all study variables are reported in Table 2.

2.3.1. Dispositional measures

The following items were assessed without reference to the particular event or Requester that the Worker chose to describe. These include the dispositional factors – i.e., those predictors that are stable across HITs – that may predict job satisfaction as a stable attitude.

2.3.1.1. Demographics. Participants completed demographics including questions about MTurk and outside work.

2.3.1.2. Personality traits. The Big Five personality traits – Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness to Experience – were assessed using two items each from the Mini-IPIP (Donnellan, Oswald, Baird, & Lucas, 2006). A sample item for Extraversion is *Am the life of the party*. Reliabilities for the Big Five traits ranged from $\alpha = .67$ to .82. Core-self evaluations ($\alpha = .77$) were assessed using the Core Self-Evaluations Scale (Judge, Erez, Bono, & Thoresen, 2003). A sample item is *I am confident I get the success I deserve in life*. Response options for both measures ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*).

2.3.1.3. Workplace friendships. Workers' friendships with other Workers ($\alpha = .90$) were assessed using the Workplace Friendship Scale (Nielsen, Jex, & Adams, 2000). A sample item is *I have the opportunity to develop close friendships with other MTurk Workers*. Response options ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*). Four other items assessed whether Workers interacted with other Workers online, face-to-face, in another setting, or not at all.

2.3.2. Narrative and event details

Each Worker was randomly assigned using Qualtrics survey software to describe either a time when he/she was *very satisfied* (i.e., Positive Event condition) or *very dissatisfied* (i.e., Negative Event condition) with the way a Requester treated him/her or the HIT he/she completed for a Requester. This item served as the basis for the critical incident approach, as a way to identify positive and negative events that related to the job satisfaction of MTurk Workers. Workers completed all remaining measures with a focus specifically on the event they described in the narrative item.

2.3.2.1. Job characteristics. Job characteristics were measured as a situational variable using two items each from the Job Diagnostic Survey (Hackman & Oldham, 1974). A sample item measuring Dealing with Others was *The HIT required a lot of cooperative work with other people*. Response options ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*). Job characteristics showed relatively low reliability (cf. Nunnally, 1978) in this study. Future research may be helpful to evaluate whether these characteristics are truly not meaningful to MTurk Workers. The microtask nature of HITs may lend to this perception. Task significance was particularly unreliable ($\alpha = .12$) so it was excluded from analyses. A sample item measuring task significance was *The HIT was one where a lot of other people could be affected by how well my work got done*.

2.3.2.2. Perceived Requester support. Perceived organizational support was measured as a situational variable called Perceived Requester support ($\alpha = .96$) using the Survey of Perceived Organizational Support – Shortened Version (Hochwarter, Kacmar, Perrewe, & Johnson, 2003). A sample item is *This Requester would have forgiven an honest mistake on my part*. Response options ranged from 1 (*Strongly disagree*) to 7 (*Strongly agree*).

2.3.2.3. Positive and negative affect. State positive ($\alpha = .93$) and negative affect ($\alpha = .88$) were measured as an interactive (i.e., dependent on both the person and situation) variable using the

Table 1
Demographics.

Variable	US	India
	M (*SD*)	
Age	32.55 (9.93)	31.62 (9.28)
Tenure on MTurk (in months)	15.86 (15.71)	35.37 (13.66)
Gender	*n* (%)	
Male	118 (52.44%)	89 (67.4%)
Female	107 (47.56%)	43 (32.6%)
Education	*n* (%)	
High school/GED	25 (11.11%)	2 (1.52%)
Some college	92 (40.89%)	6 (4.55%)
Associate's degree	3 (1.33%)	0 (.00%)
Bachelor's degree	82 (36.44%)	79 (59.85%)
Master's degree	20 (8.89%)	45 (34.09%)
Advanced graduate work or Ph.D.	3 (1.33%)	0 (.00%)
Among Workers who do not consider working on MTurk their primary job	*M* (*SD*)	
Hours worked per week outside of MTurk	35.83 (10.98)	42.19 (9.74)
Hours worked per week on MTurk	18.25 (11.12)	19.27 (12.39)
n (% of sample)	138 (61.33%)	78 (59.09%)
Among Workers who consider working on MTurk their primary job	*M* (*SD*)	
Hours worked per week outside of MTurk	20.38 (21.67)	27.77 (16.34)
Hours worked per week on MTurk	35.22 (16.65)	31.26 (23.38)
n (% of sample)	87 (38.67%)	54 (40.91%)

Positive and Negative Affective Schedule (Watson, Clark, & Tellegen, 1988). A sample item measuring negative affect is *Hostile*. Response options ranged from 1 (*Very slightly or not at all*) to 5 (*Extremely*).

2.3.2.4. Psychological contract fulfillment. Psychological contract fulfillment ($\alpha = .94$) was measured as an interactive variable using the Psychological Contract Fulfillment scale (Henderson, Wayne, Shore, Bommer, & Tetrick, 2008). A sample item from this scale is *This Requester kept his or her promises to me*. Response options ranged from 1 (*Strongly disagree*) to 7 (*Strongly agree*).

2.3.2.5. Intrinsic motivation. Intrinsic motivation ($\alpha = .74$) was measured as an interactive variable using two subscales of the Flow Dimension scale (Webster, Trevino, & Ryan, 1993). A sample item is *When I completed the HIT, I was totally absorbed in what I was doing*. Response options ranged from 1 (*Strongly disagree*) to 7 (*Strongly agree*).

2.3.2.6. Pay satisfaction. Pay satisfaction (Kuder-Richardson's $\rho{-}20 = .93$) was measured as an interactive variable using the pay satisfaction facet scale from the Job in General Satisfaction measure (Ironson, Smith, Brannick, Gibson, & Paul, 1989). A sample item is *Enough to live on*. Response options were *Yes*, *Can not decide*, and *No*.

2.3.2.7. Job satisfaction. Job satisfaction ($\alpha = .88$) was measured using the Brief Index of Affective Job Satisfaction (Thompson & Phua, 2012). A sample item is *I felt fairly well satisfied with the HIT*. Response options ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*).

2.3.2.8. Turnover. Turnover ($\alpha = .97$) was measured using four items written by the authors to assess the unique type of turnover between Workers and individual Requesters. Sample items are *I would be happy to complete more HITs for this Requester* and *If I saw the Requester post more HITs, I would NOT work on them* (reverse-scored). Response options ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*).

2.3.2.9. Sharing with coworkers. One item was used to assess whether Workers had shared information with other MTurk Workers about the experience described. Response options were *Yes* or *No*.

2.3.2.10. Situational Affordances at Work Scale (SAWS). Participants responded to a 54-item measure of 14 workplace affordances (Brawley & Pury, 2014). All items used the stem *My work environment is characterized as/by …* and response options ranged from 1 (*Not at all*) to 7 (*Very much*). A sample item for the measuring Affordances for Ownership by Others was *Someone else in charge*. Analyses relating this measure to job satisfaction and turnover are reported in another manuscript.

2.3.3. Attention-checking item

One attention-checking item asked for a particular response: *The answer to this item should be Neutral so we know to keep your data*. Response options ranged from 1 (*Strongly disagree*) to 5 (*Strongly agree*), with the correct answer being 3 (*Neutral*). Responses from Workers who did not endorse any response or endorsed any response other than 3 (*Neutral*) were excluded from analysis.

2.4. Narrative coding

Four undergraduate research assistants familiar with MTurk read the narratives to generate coding categories for positive and negative experiences based on Requester behaviors described by Workers in these experiences. The coders identified 21 categories of negative Requester behaviors and 18 categories of positive Requester behaviors. Next, two undergraduate research assistants each coded the narratives according to the categories. Initial inter-rater agreement was 89.17% for the positive narratives and 93.47% for the negative narratives. The first author served as a third coder to resolve each instance of disagreement between the raters.

3. Results

Missing data were imputed using the Expectation Maximization (EM) algorithm (Newman, 2003; Schafer & Graham, 2002). Regression results are reported in Tables 3–6.

Table 2
Descriptives.

Variable	M (SD)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.
1. Extraversion	2.84 (.94)	.77																						
2. Agreeableness	3.86 (.79)	.31	.82																					
3. Neuroticism	2.55 (.85)	−.30	−.14	.70																				
4. Openness	3.95 (.72)	.31	.38	−.17	.68																			
5. Conscientiousness	3.73 (.75)	.17	.33	−.31	.20	.67																		
6. Core self-evaluations	3.62 (.55)	.43	.29	−.63	.31	.48	.77																	
7. Coworker friendships	4.04 (1.49)	.09	.11	−.05	.04	−.05	.09	.90																
8. Skill variety	2.63 (1.04)	.10	−.06	−.01	−.02	.05	.06	−.08	.52															
9. Task significance	2.76 (.90)	.10	−.05	−.01	−.12	.02	.11	.00	.30	.12														
10. Task identity	3.75 (1.02)	−.05	−.01	.00	.07	−.07	−.02	.00	−.06	.03	.60													
11. Dealing with others	1.90 (.88)	.12	−.12	.06	−.10	.03	.03	−.05	.32	.33	−.24	.54												
12. Feedback from agents	2.93 (1.22)	−.02	.02	−.02	−.03	.04	.04	−.02	.17	.33	.22	.11	.63											
13. Feedback from the HIT	2.87 (1.06)	.08	.03	−.07	−.10	.01	.11	.06	.11	.39	.19	.19	.52	.54										
14. Autonomy	3.41 (.97)	.03	−.05	−.03	−.04	−.02	.05	.04	.14	.26	.41	−.01	.36	.38	.48									
15. Perceived requester support	3.94 (2.00)	.05	−.01	−.12	−.09	−.02	.12	.13	.16	.32	.26	.03	.50	.46	.44	.96								
16. Positive affect	3.31 (1.04)	.41	.20	−.15	.05	.20	.37	.13	.04	.22	.05	.13	.25	.28	.19	.39	.93							
17. Negative affect	1.62 (.71)	−.06	−.09	.17	−.20	−.18	−.19	.07	−.05	−.08	−.12	.02	−.22	−.08	−.16	−.33	−.19	.88						
18. Psych. contract fulfillment	4.44 (2.20)	.01	−.01	−.07	−.07	−.05	.10	.08	.16	.30	.28	.01	.47	.42	.43	.91	.30	−.36	.94					
19. Intrinsic motivation	3.54 (.78)	.20	.34	−.11	.13	.25	.30	.09	.03	.17	.08	.01	.23	.26	.16	.41	.53	−.28	.34	.74				
20. Pay satisfaction	2.17 (.74)	.03	.00	−.03	−.06	−.06	.12	.14	.20	.33	.17	.08	.41	.37	.37	.76	.34	−.25	.76	.32	.93			
21. Job satisfaction	3.32 (1.15)	.15	.12	−.08	−.04	.07	.24	.09	.13	.29	.16	.09	.37	.36	.28	.69	.61	−.31	.65	.61	.63	.88		
22. Turnover	2.60 (1.63)	−.03	.00	.08	.09	.04	−.10	−.09	−.15	−.31	−.24	−.01	−.48	−.43	−.42	−.86	−.37	.36	−.88	−.38	−.77	−.74	.97	
23. Sharing information[a]	1.39 (.49)	−.02	.10	−.08	.22	.01	.06	.35	−.03	−.10	−.07	−.04	−.08	−.06	−.03	−.09	−.06	−.03	−.07	−.04	−.06	−.11	.10	–

Note. $N = 357$. Cronbach's alpha reported on diagonal.
[a] Spearman's ρ is reported for this categorical predictor.

3.1. Predicting job satisfaction

Twelve total multiple regressions were conducted, with one each for good and bad experiences described by Indian and US MTurk Workers. Four regressions examined the dispositional, situational, and interactive factors as predictors of job satisfaction; four logistic regressions examined these factors as predictors of information sharing, a dichotomous outcome; and four regressions examined job satisfaction as a predictor of turnover. As an example, Table 3 reports the results for Indian MTurk Workers describing negative events. First, we report the raw regression coefficients for each variable predicting job satisfaction. For example, every unit increase in feedback from agents increases job satisfaction by .31 units ($SE = .13$). Task significance was excluded from analyses because it was highly unreliable (see Section 2.3.2.1).

Next in each table, we report the overall variance explained (R^2) and the change in variance explained (ΔR^2) with the addition of each set of predictors. For Indian Workers describing negative events, situational factors – i.e., characteristics of the work itself – explained the largest portion of variability in job satisfaction ($\Delta R^2 = .31$). Next, we report similar results for the prediction of information sharing. Because information sharing was measured as a dichotomous *Yes* or *No* response, we used logistic regression to analyze the same predictors for this outcome. Here, we found that dispositional variables – i.e., stable factors related to the individual, such as personality traits – predicted the largest amount of unique variance ($\Delta R^2 = .21$) in information sharing. For example, increases in coworker friendships predicted an increase in information sharing, $B = 1.20$, $SE = .57$.

Last in each table, we examined whether job satisfaction significantly predicted turnover, and found a significant negative relationship, $B = -.65$ ($SE = .14$) that predicted considerable variability in turnover from a Requester, $R^2 = .26$. This finding indicates that as job satisfaction with a specific Requester increases, Workers are less likely to quit working or refuse to work further for that Requester. Supplementary relative weight analyses, which partition the explained variance in outcomes into portions explained by each predictor (Tonidandel & LeBreton, 2011), are available upon request.

As mentioned previously, situational variables explained the most unique variance in job satisfaction for Indian Workers describing negative events ($\Delta R^2 = .31$; Table 3). However, among US MTurk Workers, interactive person-situation variables (e.g., positive affective reactions to the HIT) explained the most unique variance in job satisfaction in negative events ($\Delta R^2 = .27$; Table 4). Therefore, we partially supported *Hypothesis 1a*, that situational variables would be the strongest predictors of job satisfaction in negative work experiences on MTurk. Interactive variables also predict considerable unique variance in job satisfaction among Workers from India describing negative events ($\Delta R^2 = .24$), indicating that interactive variables may be important predictors of job satisfaction in negative events in both countries.

Dispositional variables explained the most unique variance in job satisfaction in positive events for Workers from India ($\Delta R^2 = .26$; Table 5), but interactive variables explained the most unique variance in job satisfaction with positive events for US Workers ($\Delta R^2 = .27$; Table 6) though dispositional variables did explain nearly as much unique variance ($\Delta R^2 = .24$). *Hypothesis 1b* – that personality variables would most strongly predict job satisfaction in positive events – was therefore partially supported. For crowdsourcing employers, this may mean that, when micro-tasks are designed adequately and perceived positively, Workers' individual differences will still affect their satisfaction with a particular HIT. This result may also indicate that good Requester

Table 3
Indian workers describing negative events.

Variable	Criterion			
	Job satisfaction		Sharing with coworkers	
	B (S.E.)	R^2 (ΔR^2)	*B* (S.E.)[a]	R^2_L (ΔR^2_L)
Intercept	2.09 (2.04)		−3.98 (11.46)	
Step 1: Person variables		.11		.21
Extraversion	−.28 (.20)		.91 (.96)	
Agreeableness	.14 (.24)		.65 (1.38)	
Neuroticism	−.21 (.20)		$-1.98^{\dagger}$ (1.25)	
Openness	−.32 (.25)		1.58 (1.42)	
Conscientiousness	−.22 (.21)		−.89 (1.44)	
Core self-evaluations	.00 (.43)		−4.37 (3.17)	
Coworker friendships	$-.15^{\dagger}$ (.08)		1.20^{*} (.57)	
$F(7, 56) = .95$				
Step 2: Situation variables		.42 (.31)		.40 (.18)
Skill variety	.00 (.13)		−.20 (.68)	
Task significance[b]	n/a		n/a	
Dealing with others	−.24 (.15)		1.27 (.92)	
Feedback from agents	$.31^{*}$ (.13)		$.96^{\dagger}$ (.61)	
Autonomy	−.06 (.18)		.27 (.94)	
Feedback from HIT	−.20 (.15)		−.71 (.86)	
Task identity	.12 (.12)		−.19 (.62)	
Perceived requester support	$.25^{\dagger}$ (.13)		-1.72^{*} (.82)	
$F(14, 49) = 2.54^{**}$				
Step 3: Interactive variables		.66 (.24)		.47 (.07)
Negative affect	−.18 (.18)		1.37 (1.13)	
Positive affect	.26† (.15)		1.38† (.81)	
Psychological contract fulfillment	.10 (.14)		.66 (.72)	
Intrinsic motivation	$.67^{**}$ (.24)		.99 (1.48)	
Pay satisfaction	.42† (.21)		−.77 (1.06)	
$F(19, 44) = 4.58^{***}$				

Variable	Criterion: Turnover	
	B (S.E.)	R^2
Intercept	5.37 (.48)	.26
Job satisfaction	$-.65^{***}$ (.14)	
$F(1, 62) = 21.32^{***}$		

Note. $n = 64$. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.
[a] These parameter estimates are in logit form.
[b] Task significance was highly unreliable (see Table 2) and not analyzed.

practices on MTurk result in a “weaker” workplace where situational cues for satisfaction responses are relatively weak and allow for more personality-relevant variability in behavior and attitudes (see Meyer & Dalal, 2009).

Intrinsic motivation was significantly positively related to job satisfaction in both narrative conditions for both cultural groups, while four other variables were significantly related to job satisfaction in only a few conditions or for only one group. Feedback from the HIT itself was significantly positively related to job satisfaction in positive events for Indian Workers. Agreeableness and state positive affect were positively related to job satisfaction in positive events for US Workers, while only state positive affect was positively related to job satisfaction in negative events for US Workers.

3.2. Predicting turnover

In both narrative conditions for both cultural groups, job satisfaction was consistently negatively related to turnover, supporting *Hypothesis 2*. The traditional relationship between job satisfaction and turnover appears to hold true in the context of working on MTurk, with job satisfaction explaining 11 to 26 percent of the variability in turnover. As discussed previously, MTurk – and crowdsourcing sites in general – provide the employee with relatively increased freedom and reduced cost to quit when one is dissatisfied with a particular Requester, emphasizing the importance of studying job satisfaction and other predictors of turnover among crowdsourced employees.

3.3. Predicting information sharing

Sharing information about Requesters with other MTurk Workers was related to the friendships a Worker has with other Workers in both narrative conditions for both cultural groups, and seven additional variables predicted information sharing in at least one condition for one group (*Research Questions 1 and 2*). Perceived Requester support negatively predicted Indian Workers sharing information about negative experiences, while feedback from agents negatively predicted US Workers sharing information about negative experiences. For US Workers only, openness and dealing with others positively predicted sharing information about positive experiences; autonomy and pay satisfaction negatively predicted sharing about positive experiences; conscientiousness and feedback from agents negatively predicted sharing about negative experiences; and autonomy positively predicted sharing about negative experiences.

Requesters may benefit from being aware of the network of Workers and their information sharing behaviors and consistently

Table 4
US workers describing negative events.

Variable	Criterion: Job satisfaction B (S.E.)	R^2 (ΔR^2)	Sharing with coworkers B (S.E.)[a]	R^2_L (ΔR^2_L)
Intercept	.41 (1.17)		1.98 (4.20)	
Step 1: Person variables		.12		.18
Extraversion	−.19 (.11)		−.61 (.38)	
Agreeableness	−.05 (.11)		.23 (.43)	
Neuroticism	−.01 (.14)		−.76 (.48)	
Openness	−.14 (.14)		.86† (.48)	
Conscientiousness	−.14 (.13)		−1.49** (.55)	
Core self-evaluations	.32 (.22)		−.23 (.83)	
Coworker friendships	.03 (.06)		1.17*** (.31)	
$F(7, 99) = 1.90^{\dagger}$				
Step 2: Situation variables		.22 (.10)		.36 (.18)
Skill variety	.02 (.09)		.26 (.33)	
Task significance[b]	n/a		n/a	
Dealing with others	.09 (.13)		−.11 (.43)	
Feedback from agents	−.06 (.08)		−1.05*** (.32)	
Autonomy	−.13 (.09)		.74* (.36)	
Feedback from HIT	−.02 (.11)		.02 (.36)	
Task identity	.06 (.08)		−.56† (.30)	
Perceived requester support	.18 (.14)		−.99† (.53)	
$F(14, 92) = 1.81^{*}$				
Step 3: Interactive variables		.49 (.27)		.39 (.03)
Negative affect	−.18 (.12)		.75† (.44)	
Positive affect	.30** (.11)		.43 (.39)	
Psychological contract fulfillment	.03 (.10)		.04 (.34)	
Intrinsic motivation	.52*** (.11)		−.02 (.39)	
Pay satisfaction	.08 (.15)		−.01 (.49)	
$F(19, 87) = 4.32^{***}$				

Variable	Criterion: Turnover B (S.E.)	R^2
Intercept	5.42 (.21)	.22
Job satisfaction	−.47*** (.09)	
$F(1, 105) = 30.33^{***}$		

Note. $n = 107$. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.
[a] These parameter estimates are in logit form.
[b] Task significance was highly unreliable (see Table 2) and not analyzed.

making decisions that positively affect one's reputation among Workers. Results of this study indicate that friendships with other Workers consistently predict information sharing across cultures and types of experiences. Post hoc analyses of the four items assessing Worker interactions with other Workers indicated that a majority of Workers (55%) interact with other Workers online, while 21% interact with other Workers in person, and 5% interact with other Workers in another setting. Less than half of Workers — 39% — reported not interacting with other Workers at all. Therefore, the majority of Workers may be likely to share information about their experience with a Requester with other Workers.

However, reports of sharing information alone do not necessarily detail what type of information is shared: it is possible that Workers use information sharing to fill a need to connect with other Workers in addition to gathering work-related information (e.g., Salehi et al., 2015; Schmidt, 2015). While specific content shared was not studied here, responses in our data indicated that over one-third of Workers (37%) shared information with other Workers in good experiences, and a slightly higher 41% of Workers shared information with other Workers in bad experiences. Therefore, Requesters may find that at least some Workers share information with other Workers about their experiences (see also Bederson & Quinn, 2011; Marshall & Shipman, 2013; Martin et al., 2014). However, these results also indicate that information sharing about negative events may be reduced by positive Requester behaviors that promote perceptions of Requester support and feedback, as well as Worker autonomy.

3.4. Positive and negative crowdsourcing employer behaviors

The job satisfaction ratings associated with each of the positive and negative Requester behaviors identified in the narratives are reported in Table 7 (*Research Question 3*). Sample narrative answers for each category are provided in the Appendix. Shown in the table, the positive Requester behaviors resulting in the highest five average job satisfaction ratings were: building an ongoing relationship with Workers, providing encouraging feedback, posting simple HITs that paid well, posting HITs that were interesting, and providing a progress bar. We note that only two Workers described including a progress bar in their positive experience, so this result should be interpreted with caution. However, this feature is typically simple to implement so we do recommend including one to foster positive Worker reactions.

The negative Requester behaviors resulting in the lowest five average job satisfaction ratings were: paying an unfair wage, including difficult attention check questions, using majority rules (essentially, inter-Worker agreement) for rejection decisions, advertising a HIT as taking less time than it actually does, and

Table 5
Indian workers describing positive events.

Variable	Criterion: Job satisfaction B (S.E.)	Job satisfaction R^2 (ΔR^2)	Sharing with coworkers B (S.E.)[a]	Sharing with coworkers R^2_L (ΔR^2_L)
Intercept	.55 (1.19)		−5.21 (7.12)	
Step 1: Person variables		.26		.13
Extraversion	−.02 (.09)		.14 (.54)	
Agreeableness	−.07 (.10)		−.26 (.61)	
Neuroticism	.06 (.10)		.10 (.65)	
Openness	.07 (.12)		.28 (.71)	
Conscientiousness	.11 (.12)		1.07 (.72)	
Core self-evaluations	.17 (.23)		−2.19 (1.49)	
Coworker friendships	.02 (.04)		.63*** (.25)	
Step 2: Situation variables		.33 (.07)		.23 (.10)
Skill variety	−.06 (.06)		.03 (.40)	
Task significance[b]	n/a		n/a	
Dealing with others	−.09 (.07)		.02 (.42)	
Feedback from agents	−.05 (.09)		.36 (.51)	
Autonomy	−.15 (.09)		.54 (.52)	
Feedback from HIT	.20* (.08)		.14 (.49)	
Task identity	−.01 (.09)		−.20 (.52)	
Perceived requester support	−.05 (.10)		1.17† (.65)	
Step 3: Interactive variables		.54 (.21)		.25 (.01)
Negative affect	.12 (.13)		−.27 (.79)	
Positive affect	.27†(.14)		.43 (.88)	
Psychological contract fulfillment	.00 (.08)		−.04 (.50)	
Intrinsic motivation	.45** (.15)		−.80 (.99)	
Pay satisfaction	.06 (.21)		−.97 (1.32)	

Variable	Criterion: Turnover B (S.E.)	R^2
Intercept	3.82 (.53)	.24
Job satisfaction	−.56*** (.12)	
$F(1, 66) = 20.73$***		

Note. $n = 68$. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.
[a] These parameter estimates are in logit form.
[b] Task significance was highly unreliable (see Table 2) and not analyzed.

blaming others (e.g., other members of one's research team) for rejection and seemingly refusing to investigate further or consult that person. We recommend that crowdsourcing employers use these findings – both positive and negative – as guidelines when designing HITs, deciding on compensation, and interacting with Workers.

Fair pay on MTurk has been discussed in detail by many researchers (see Section 1.1). One recommended practice is to pilot a HIT on MTurk and set pay based on the observed HIT submission time, but our evidence supports other researchers' caution to Requesters in using HIT or survey completion time alone to determine pay (cf. Marshall & Shipman, 2013). Based on MTurk job management data, Workers accepted and submitted our HIT in an average of 39 min ($SD = 86$ min). However, this metric includes all participants without exclusion as was done for analysis. This estimate also appears to include individuals who did not complete the HIT immediately (maximum submission time = 21.50 h) or who completed the task prior to accepting, entering the completion code, and submitting the HIT (minimum submission time = 4 s). Based on Qualtrics start and end times for the final sample of participants included in analysis, we observed a shorter average completion time of 30 min, but the observed variability was still large, $SD = 28$ min. The minimum completion time observed with this method was 6 min, and the maximum was 4 h and 16 min.

Based on these observations, we caution users against using recorded completion times alone to determine pay, especially average completion times that may include inattentive responses. Median completion times may provide a more moderate estimate, with less influence by Workers who completed the task before accepting and submitting the HIT or Workers who held the HIT before submitting, fearing rejection based on quick work, or waited to complete the HIT. The median completion time recorded with the second method based on the final sample – 22 min – was closest to our estimated completion time of 15–20 min, and resulted in a median pay approximately 1.4 times the federal minimum hourly wage.

4. Discussion

The present study is the first examination of MTurk Workers' job satisfaction and turnover. This study suggests that intrinsic motivation is a strong predictor of job satisfaction among crowdsourced employees from both the US and India in both positive and negative experiences. Furthermore, we established that feedback may be a stronger predictor of job satisfaction among Indian Workers, while trait agreeableness and state positive affectivity may be stronger predictors of job satisfaction among US Workers. Overall, we found that dispositional factors predicted considerable unique variance in job satisfaction in positive experiences, while situational factors predicted the most variance in job satisfaction for Indian Workers in negative experiences and person-situation interactive factors predicted the most variance in job satisfaction for US Workers in negative experiences. Job satisfaction in turn was consistently related to Requester-specific turnover.

Table 6
US workers describing positive events.

Variable	Criterion			
	Job satisfaction		Sharing with coworkers	
	B (S.E.)	R^2 (ΔR^2)	B (S.E.)[a]	R^2_L(ΔR^2_L)
Intercept	.14 (.75)		−9.46 (5.37)	
Step 1: Person variables		.24		.25
Extraversion	.00 (.05)		−.60† (.33)	
Agreeableness	.18** (.06)		−.26 (.44)	
Neuroticism	.03 (.06)		.57 (.40)	
Openness	.01 (.06)		1.54** (.52)	
Conscientiousness	−.07 (.06)		.29 (.42)	
Core self-evaluations	.03 (.11)		.67 (.77)	
Coworker friendships	−.05 (.03)		1.29*** (.30)	
F(7, 110) = 4.89***				
Step 2: Situation variables		.29 (.05)		.32 (.07)
Skill variety	.04 (.04)		−.11 (.28)	
Task significance[b]	n/a		n/a	
Dealing with others	.07 (.06)		.87* (.42)	
Feedback from agents	−.08† (.04)		.28 (.32)	
Autonomy	.00 (.05)		−.82* (.38)	
Feedback from HIT	.05 (.05)		.02 (.35)	
Task identity	.07 (.06)		−.35 (.38)	
Perceived requester support	−.01 (.05)		−.37 (.38)	
F(14, 103) = 2.96**				
Step 3: Interactive variables		.55 (.27)		.39 (.07)
Negative affect	.03 (.09)		−.36 (.77)	
Positive affect	.25*** (.06)		−.54 (.39)	
Psychological contract fulfillment	.10 (.07)		.59 (.54)	
Intrinsic motivation	.22** (.07)		.94† (.53)	
Pay satisfaction	.24† (.13)		−2.25* (.99)	
F(19, 98) = 6.38***				

Variable	Criterion: Turnover	
	B (S.E.)	R^2
Intercept	2.37 (.30)	.11
Job satisfaction	−.28*** (.08)	
F(1, 116) = 13.62***		

Note. $n = 118$. $^†p < .10$, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.
[a] These parameter estimates are in logit form.
[b] Task significance was highly unreliable (see Table 2) and not analyzed.

4.1. Theoretical recommendations

The results of the present study suggest that at least some of the predictors of job satisfaction in traditional jobs are relevant to crowdsourced workers, and the relationship between job satisfaction and turnover was consistently supported in this new work context. However, the majority of the predictors did not individually significantly predict job satisfaction and information sharing, perhaps indicating that entirely different variables or different operationalizations of current variables are required to better understand crowdsourced workers' attitudes and behaviors. Task significance — a traditional job characteristic — in particular was unreliable in the present study. Although it was assessed using only two items, this finding may preliminarily show that the broader impact of one's work might not be a relevant concept in the domain of crowdsourced microtasks.

Even though some of the observed relationships were comparable to more traditional I–O psychology findings, the operational definitions of the variables differed considerably from previous research. For example, in order to study turnover in this study, we needed to narrowly define it as Requester-specific turnover (rather than turnover from MTurk as a whole) and write items that were reasonable for the crowdsourcing setting, referencing the Worker's decision to stop working for a Requester and refuse to future opportunities to work for a Requester. While this represents an interesting opportunity to study turnover behaviors instead of turnover intentions, it alters the way in which we can study the concept. For example, where it may be impractical to study turnover in a traditional organization, it may be equally impractical to study turnover intentions in a sample of crowdsourced employees, who may simply quit immediately as desired.

Other researchers have suggested that traditional work theories can be adapted to the crowdsourced work context. For example, performance-based pay has been recommended for tasks where Workers can increase work quality by increasing time spent (Ho et al., 2015), which resembles Expectancy from Vroom's theory of work motivation (1964). The demands-abilities component of person-job fit theory may also apply similarly, while the needs-supplies fit component may require new concepts of payment fit, enjoyment fit, and time fit (Schulze et al., 2012).

The changing nature of work may impact the meaning of and worker expectations regarding many other constructs, such as procedural, distributive, and interactional justice (Ryan & Wessel, 2015). Overall, I–O psychology and psychology in general must remain flexible in its study of work attitudes and behaviors in order to remain relevant to the changing nature of work, which now includes crowdsourced work on platforms like MTurk.

Table 7
Positive and negative requester behaviors described by workers.

	# (%)	Job satisfaction *M* (*SD*)	
		Requester did this behavior	Requester did not do this behavior
Positive behaviors			
Building relationship	19 (10.22%)	4.38 (.54)	3.95 (.58)
Encouraging feedback	13 (6.99%)	4.33 (.40)	3.97 (.59)
Simple HIT that paid well	17 (9.14%)	4.32 (.56)	3.96 (.58)
Interesting HIT	34 (18.28%)	4.27 (.52)	3.93 (.59)
Progress bar	2 (1.08%)	4.13 (.88)	3.99 (.59)
Receiving bonus for good work	54 (29.03%)	4.08 (.54)	3.96 (.61)
Approved HIT quickly	22 (11.83%)	4.06 (.53)	3.99 (.60)
Fair pay	52 (27.96%)	4.00 (.49)	3.99 (.63)
Pay matches work/effort	17 (9.14%)	3.99 (.40)	4.00 (.60)
Professional communication	25 (13.44%)	3.97 (.53)	4.00 (.60)
Gives good instructions	8 (4.30%)	3.91 (.38)	4.00 (.60)
Receptive to feedback	18 (9.68%)	3.88 (.76)	4.01 (.57)
Quick response	47 (25.27%)	3.84 (.56)	4.05 (.59)
Forgiving technical issues	29 (15.59%)	3.77 (.54)	4.04 (.59)
Forgiving worker mistake	40 (21.51%)	3.76 (.62)	4.06 (.57)
Rejection reversed after issue	10 (5.38%)	3.68 (.59)	4.01 (.59)
Forgive missing completion code	5 (2.69%)	3.50 (.40)	4.01 (.59)
Adequate time	3 (1.61%)	3.33 (.76)	4.01 (.58)
Overall positive experiences		4.00 (.59)	
Negative behaviors			
Unfair pay	17 (9.94%)	1.53 (.74)	2.70 (1.15)
Difficult attention check questions	3 (1.75%)	1.67 (.63)	2.60 (1.17)
Majority rules for rejection decisions	3 (1.75%)	1.75 (.50)	2.60 (1.17)
HIT took longer than advertised	15 (8.77%)	1.78 (.94)	2.66 (1.16)
Outsourcing blame for rejections	2 (1.17%)	2.00 (.00)	2.59 (1.17)
Giving negative feedback	4 (2.34%)	2.13 (1.79)	2.60 (1.15)
Mass reject/taking advantage of Workers	34 (19.88%)	2.40 (1.03)	2.63 (1.20)
Technical difficulties caused rejection	23 (13.45%)	2.41 (1.23)	2.61 (1.16)
Insufficient time allowed for HIT	10 (5.85%)	2.43 (1.45)	2.60 (1.15)
Unwilling to compromise	14 (8.19%)	2.46 (.76)	2.60 (1.20)
No or bad completion code	15 (8.77%)	2.48 (1.12)	2.60 (1.17)
Rejected good work	90 (52.63%)	2.51 (1.12)	2.67 (1.22)
Rejected for “working too quickly”	8 (4.68%)	2.56 (.84)	2.59 (1.18)
Not responding	55 (32.16%)	2.59 (1.15)	2.58 (1.18)
No rejection reason provided	34 (19.88%)	2.63 (1.14)	2.58 (1.17)
Accidentally rejecting HITs	3 (1.75%)	2.75 (1.09)	2.58 (1.17)
Hurting ratings	14 (8.19%)	2.93 (1.17)	2.55 (1.16)
Blocking for no reason	21 (12.28%)	2.96 (1.20)	2.53 (1.16)
Long delay in pay	4 (2.34%)	3.06 (1.48)	2.57 (1.16)
Revealing qualification late	7 (4.09%)	3.11 (1.49)	2.56 (1.15)
Unclear instructions	28 (16.37%)	3.12 (1.05)	2.48 (1.16)
Overall negative experiences		2.59 (1.17)	

Note. See Appendix for examples of each category.

4.2. Practical recommendations

Because our study indicates that approximately 40% of MTurk Workers consider working on MTurk their primary job (see Table 1), we strongly encourage individuals who use MTurk for research purposes to act as reputable employers. This finding exceeds reports from 2009 that up to 14% of US Workers and 27% of Indian Workers sometimes or always require income from MTurk to make ends meet (Ross et al., 2010), as well as reports from 2010 that 10% of US Workers and 25% of Indian Workers use income from MTurk for primary needs such as groceries (Ipeirotis, 2010b). Our increased finding may reflect an error of sampling difference, a true change in response to different ways of asking a similar question, or a true change over time. We must also consider the fact that this response was self-reported and not authenticated by the MTurk platform in any way.

Ultimately, we encourage Requesters making decisions regarding the pay of research participants on MTurk to consider the fact that some non-negligible subset of one's own Workers – perhaps between 10 and 40 percent – may be relying on MTurk as a primary job or for necessary income. As noted in Section 3.4, unfair wages and inaccurately listed time requirements were among the top five worst Requester behaviors, resulting in the lowest average Worker job satisfaction. We have used the US federal minimum wage (i.e., $7.25 for an hour of work) as a standard for determining pay, though we caution Requesters against the use of average completion time alone to determine the amount of time required by a task (see Section 3.4).

We provide several evidence-based recommendations for crowdsourcing employers based on our quantitative and qualitative findings. First, whenever possible, crowdsourcing employers should develop intrinsically motivating or interesting tasks. Second, crowdsourcing employers should budget their tasks to pay fairly. Third, crowdsourcing employers should implement positive communication practices with Workers, such as providing helpful and timely feedback and addressing Worker concerns. The results of this study indicate that crowdsourcing work is not unlike traditional work in some ways – intrinsic motivation, feedback, trait agreeableness and state positive affect all affect job satisfaction; job satisfaction is negatively related to turnover – but may differ from traditional work in other ways. For example, task significance may not be a meaningful construct in this context, and shared information about employers has novel outlets and prominence compared to traditional work (but see glassdoor.com for one notable example for traditional employees).

4.3. Limitations and future research

The present study provides an initial investigation of the variables that predict job satisfaction, turnover, and sharing information among crowdsourced workers in particularly good and bad events. Future research on less extreme incidents may allow us to understand more about typical crowdsourced work experiences. Future research on other antecedents of job satisfaction and turnover – such as multiple job holding – as well as outcomes – such as strain – would be helpful in understanding the broader impact and long-term outcomes of crowdsourcing work.

One limitation of the present study is the use of self-reported data from single sources at a single time point, which raise concerns about common method variance and other biases in the data. While future studies should use other research designs where possible, we note that, conversely, this design precludes other likely issues in studying this topic: first, gathering meaningful reports from multiple sources about an individual MTurk work experience may be impossible; second, splitting data collection into multiple sessions may introduce concerns about suggestible or exaggerated memories of events (e.g., Loftus, Miller, & Burns, 1978).

Researchers studying the work experiences of crowdsourced employees may benefit from focusing on the variables identified here, either as classes of variables (e.g., dispositional) or as discrete predictors (e.g., intrinsic motivation) as predictors of satisfaction with microtask work. Researchers may also study the work experiences of crowdsourced employees by focusing on the interaction between Workers and Requesters; for example, results of this study suggest that feedback from agents mitigates negative information sharing by US Workers, so researchers could investigate specific ways to improve Workers' perceptions of the quality of feedback from Requesters.

In future studies of work experiences on MTurk, researchers should ensure that constructs and measures are adapted appropriately for studying work on MTurk. Here, for example, we adapted all references to one's job to refer to the specific HIT that the Worker described in their narrative response. Many other commonly used surveys may require review before further use in this new context. As an example, the measure used to assess pay satisfaction in the present study, from the Job in General Satisfaction measure (Ironson et al., 1989), includes the item, *Enough to live on*. Such items may inaccurately assess satisfaction with pay from MTurk if one operates under the assumption that MTurk is or is not likely to provide many Workers with full-time or otherwise important employment, as discussed in Section 4.2. Significant adaptation of measures and likely the underlying constructs may be required to study work on MTurk.

This study also provided a cross-cultural investigation of this work context, with a specific focus on the two countries that comprise the majority of MTurk Workers, and several differences in results were observed across these two countries. However, this study was conducted in American English, which may have impacted the results. For example, a relatively large number of Workers from India ($n = 76$, compared to $n = 5$ Workers from US) incorrectly answered the attention-checking item; we speculate that this results at least partially from language differences, which could have compromised responses to other survey items and justified our exclusion of these individuals' responses from analysis (see Section 2.3.3). Future studies should investigate the work experiences of MTurk Workers from other countries in native languages when possible. Last, MTurk is only one of many crowdsourcing sites. Future research should examine other work experiences on other crowdsourcing platforms (see Aguinis & Lawal, 2013; Teodoro, Ozturk, Naaman, Mason, & Lindqvist, 2014; Vakharia & Lease, 2015).

4.4. Conclusion

The present study demonstrated that some constructs, relationships, and recommendations from traditional work settings might be applicable to the novel work experiences of crowdsourced workers. However, our findings also indicate that at least some of what we know about the psychology of traditional work experiences may need to be adapted or developed anew in order to study these novel work experiences that are both enabled and impacted by technology. MTurk is one of a number of crowdsourcing platforms and represents one of many novel work experiences as well as broader trends in changes to work and the workforce. In order to remain applicable as work arrangements change with technology, I–O psychology and psychology in general must remain flexible in how we conceptualize and study work experiences.

Acknowledgments

This research was funded from the generous support of Clemson Creative Inquiry. We thank Dr. Matt Lease for his helpful comments on an earlier version of this manuscript. We thank Erika Fosu, Michelle Flynn, Cameron Lemere, and Briana Smith for assistance with data analysis.

Appendix

Good behaviors – sample narratives

Building relationship: “I had an MTurk task where I would complete a specific HIT weekly with a Requester about a year ago. I would contact the Requester several times over a span of four months and built a pretty positive relationship with them.”

Encouraging feedback: “There are times the Requester puts in a ‘Thank you’ for working on the HITs which is a very good gesture which sometimes makes one feel good and gives the feeling that the efforts are well appreciated.”

Simple HIT that paid well: “There was someone who was putting [item] label hits up. To simply label … with a word and submit. These were around 10 cents a pop. I spent a whole summer doing these. It was awesome.”

Interesting HIT: “I … really enjoyed the work because it appealed to my intellectual side and didn't make me feel like I was doing menial labor for an anonymous company.”

Progress bar: “There's one Requester, she posts surveys … They have an accurate measurement of progress.”

Receiving bonus for good work: “I received a bonus too. That really made it sweet for me. It made me feel good to know that I was appreciated, even though the payment was minimal.”

Approved HIT quickly: “There are numerous people that have similar HITs but I always search the same Requester because he pays me almost immediately.”

Fair pay: “I did a HIT for a Requester for \$.50. I was surprised to receive an email from him with a bonus of \$2.00 stating that because it took me longer to do the HIT than advertised, he was giving me enough to make up for that time to make sure that I received state minimum wage.”

Pay matches work/effort: “… they pay a fair wage for the amount of time and effort the task takes to complete.”

Professional communication: “Instead of rejecting my work, the Requester contacted me and told me about the error. The Requester said that he/she verified my work through my Worker ID. The Requester approved my work, and told me to be more careful in the future.”

Gives good instructions: “They even had a in-depth tutorial to make sure you knew how to work it before you accepted your first HIT.”

Receptive to feedback: “One fairly new Requester … actually came onto one of the MTurk forums and interacted with the Workers. They were trying to get a feel for how to improve the HIT.”

Quick response: “They got back to me right away which is not all that common and suggested a couple of things.”

Forgiving technical issues: “In one instance a survey did not submit properly. I notified the Requester … She was very quick to respond, and her e-mails were ‘human.’ She acknowledged the problem with some frustration and good humor, and e-mailed updates on what was happening without being prompted.”

Forgiving Worker mistake: “I did another HIT for the same Requester and failed to complete the entire HIT, due to me overlooking it, and was faced with a rejection. I messaged the Requester and he gave me a chance to do the HIT again in order to reverse my rejection.”

Rejection reversed after Worker reported problem: “A week ago, a Requester rejected my work mistakenly due to an error in their software. … I sent an email to the Requester and he looked into it right away. He said it was a mistake and overturned it quickly, and gave me a $1.00 bonus for the trouble.”

Forgive missing completion code: “I accidentally forgot to put the completion code, but it was obvious that I completed the HIT by comments I put in the comments section. The Requester checked and I was awarded compensation for the HIT.”

Adequate time: “He also was completely honest about the time it would take to do the study.”

Bad behaviors — sample narratives

Unfair pay: “Once I answered a survey that was bubble hell. There were pages upon pages of questions. There was also some writing. The survey said it would take ten minutes but it took half an hour. The pay was poor. A dollar for ten minutes is great, a dollar for half an hour is not.”

Difficult attention check questions: “I didn't summarize the experiment with the wording she wanted, and was rejected.”

Majority rules for rejection decisions: “I really, really hate when a Requester does majority rules, especially if they allow just about anyone to do their HITs. … I've been rejected because the majority of morons selected something different than me, even though I know I was right.”

HIT took longer than advertised: “There was quite a bit of writing that came up that wasn't mentioned beforehand. At one point I thought the survey started over because I was answering questions that I swore I'd answered earlier. … It took more than 30 min and paid one dollar.”

Outsourcing blame for rejections (e.g., blaming research team): “I emailed the Requester … the reply was ‘This is not my fault; my team did this.’”

Giving negative feedback: “I, and others, received rejections with unnecessarily rude comments included. It's fine to reject the work if it is unsatisfactory, but there is no need to be so mean …”

Mass reject/Requester apparently taking advantage of Workers: “I received no response from the Requester which leads me to believe that there probably wasn't anything wrong with what I wrote, but that they simply did not want to pay. I now watch out for that particular Requester … .”

Technical difficulties resulted in rejection: “At the end of the survey I ended up receiving some sort of error. I contacted the Requester and all they said was there was nothing they could do …”

Insufficient time allowed for HIT: “I got timed out of a very long and tedious survey …”

Unwilling to compromise: “I tried to contact the Requester but they were really unwilling to move and pretty much said I was just out of luck on a reversal and gave me a pretty generic explanation for why they were rejected.”

No or bad completion code: “At the end of the survey there was no survey completion code, the survey simply went back to the first page. … he rejected my HIT with no explanation given.”

Rejected good work: “I carefully read all the instructions and attention checks. … they never gave me a reason why it was rejected.”

Rejected for “working too quickly”: “I did a HIT and I did not accept it before I did it. After I did the HIT I accepted and inputted the code. I got rejected by the Requester, who stated that I could not have been paying attention to have completed it in such a short time.”

Not responding: “I emailed them about it but never heard back from them.”

No rejection reason provided: “I received a rejection on a HIT. And I wasn't sure why I received it … I contacted the Requester and he/she did not respond. They never gave me a reason why …”

Accidentally rejecting HITs: “Their computer program had a glitch.”

Hurting ratings: “Many of the surveys look very similar so it's hard to remember if I had done one before. I think there should be a better system on this, because rejections will negatively affect your overall score.”

Blocking for no reason: “Someone decided to block me from doing any more of their surveys but provided me with no information as to why.”

Long delay in pay: “The Requester indicated that the HITs would be reviewed within ‘a week's time.’ Two weeks later, those HITs are still ‘pending’ …”

Revealing qualification after accept/submit HIT: “… the Requester does not identify the job location. When a Worker accepts a HIT and goes ahead with reading the rules, the Worker comes to know that the location is not where the Worker is. Therefore, he has to return the HIT.”

Unclear instructions: “I judged the best I could but the Requester rejected my work, saying that I did not understand the instructions. There were barely any instructions to begin with.”

References

Aguinis, H., & Lawal, S. O. (2013). eLancing: a review and research agenda for bridging the science-practice gap. *Human Resource Management Review, 23*, 6–17. http://dx.doi.org/10.1016/j.hrmr.2012.06.003.

Antin, J., & Shaw, A. (2012). *Social desirability bias and self-reports of motivation: a study of Amazon Mechanical Turk in the US and India*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX.

Barger, P., Behrend, T. S., Sharek, D. J., & Sinar, E. F. (2011). IO and the crowd: frequently asked questions about using Mechanical Turk for research. *The Industrial-Organizational Psychologist, 49*(2), 11–17.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323.

Bederson, B. B., & Quinn, A. J. (2011). *Web workers unite! addressing challenges of online laborers*. Paper presented at the CHI'11 Extended Abstracts on Human Factors in Computing Systems, Vancouver, BC, Canada.

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods, 43*(3), 800–813. http://dx.doi.org/10.3758/s13428-011-0081-0.

Brawley, A. M., & Pury, C. L. S. (2014). De-balkanization: a molar organizational climate framework based on situational affordances. In *Poster presented at the 29th annual meeting of the Society of Industrial and Organizational Psychology, Honolulu, HI*.

Brown, S. P., & Peterson, R. A. (1993). Antecedents and consequences of salesperson job satisfaction: meta-analysis and assessment of causal effects. *Journal of Marketing Research, 30*, 63–77. http://dx.doi.org/10.2307/3172514.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3–5. http://dx.doi.org/10.1177/1745691610393980.

Cartwright, S. (2003). New forms of work organization: Issues and challenges. *Leadership & Organization Development Journal, 24*(3), 121–122.

Cascio, W. R., & Aguinis, H. (2011). *Applied psychology in human resource management* (7th ed.). Upper Saddle River, NJ: Pearson.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior, 29*, 2156–2160. http://dx.doi.org/10.1016/j.chb.2013.05.009.

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*, 112–130. http://dx.doi.org/10.3758/s13428-013-0365-7.

Chandler, J., Paolacci, G., & Mueller, P. (2013). Risks and rewards of crowdsourcing marketplaces. In J. Chandler, G. Paolacci, & P. Mueller (Eds.), *Handbook of human computation* (pp. 377–392). New York: Springer.

Cooper, W. H., & Withey, M. J. (2009). The strong situation hypothesis. *Personality and Social Psychology Review, 13*, 62–72. http://dx.doi.org/10.1177/1088868308329378.

Cropanzano, R., James, K., & Konovsky, M. A. (1993). Dispositional affectivity as a predictor of work attitudes and job performance. *Journal of Organizational Behavior, 14*, 595–606. http://dx.doi.org/10.1002/job.4030140609.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment, 18*, 192–203. http://dx.doi.org/10.1037/1040-3590.18.2.192.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2012). Data collection in a flat world: the strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213–224. http://dx.doi.org/10.1002/bdm.1753.

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: update, moderator tests, and research implications for the next millenium. *Journal of Management, 26*, 463–488. http://dx.doi.org/10.1177/014920630002600305.

Gupta, N., Crabtree, A., Rodden, T., Martin, D., & O'Neill, J. (2014). *Understanding Indian crowdworkers*. Paper presented at the 17th conference on Computer Supported Cooperative Work, Baltimore, MD.

Hackman, J. R., & Oldham, G. R. (1974). *The job diagnostic survey: an instrument for the diagnosis of jobs and the evaluation of job redesign projects: DTIC document*.

Henderson, D. J., Wayne, S. J., Shore, L. M., Bommer, W. H., & Tetrick, L. E. (2008). Leader–member exchange, differentiation, and psychological contract fulfillment: a multilevel examination. *Journal of Applied Psychology, 93*, 1208–1219. http://dx.doi.org/10.1037/a0012678.

Hochwarter, W. A., Kacmar, C., Perrewe, P. L., & Johnson, D. (2003). Perceived organizational support as a mediator of the relationship between politics perceptions and work outcomes. *Journal of Vocational Behavior, 63*, 438–456. http://dx.doi.org/10.1016/S0001-8791(02)00048-9.

Holden, C. J., Dennie, T., & Hicks, A. D. (2013). Assessing the reliability of the M5-120 on Amazon's Mechanical Turk. *Computers in Human Behavior, 29*, 1749–1754. http://dx.doi.org/10.1016/j.chb.2013.02.020.

Horton, J. J. (2011). The condition of the turking class: are online employers fair and honest? *Economics Letters, 111*, 10–12. http://dx.doi.org/10.1016/j.econlet.2010.12.007.

Ho, C. J., Slivkins, A., Suri, S., & Vaughan, J. W. (2015). *Incentivizing high quality crowdwork*. Paper presented at the Proceedings of the 24th International Conference on World Wide Web.

Ipeirotis, P. G. (2010a). Analyzing the Amazon Mechanical Turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students, 17*(2), 16–21. http://dx.doi.org/10.1145/1869086.1869094.

Ipeirotis, P. G. (2010b). *Demographics of mechanical turk*. CeDER Working Papers. New York University. Retrieved from http://hdl.handle.net/2451/29585.

Irani, L. C. (2015a). The cultural work of microwork. *New Media & Society, 17*, 720–739. http://dx.doi.org/10.1177/1461444813511926.

Irani, L. C. (2015b). Difference and dependence among digitial workers: the case of Amazon Mechanical Turk. *South Atlantic Quarterly, 114*, 225–234. http://dx.doi.org/10.1215/00382876-2831665.

Irani, L. C., & Silberman, M. S. (2013). *Turkopticon: Interrupting worker invisibility in Amazon Mechanical Turk*. http://dx.doi.org/10.1145/2470654.2470742. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a job in general scale: a comparison of global, composite and specific measures. *Journal of Applied Psychology, 74*, 193–200. http://dx.doi.org/10.1037/0021-9010.74.2.193.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluations scale: development of a measure. *Personnel Psychology, 56*, 303–331. http://dx.doi.org/10.1111/j.1744-6570.2003.tb00152.x.

Judge, T. A., Heller, D., & Klinger, R. (2008). The dispositional sources of job satisfaction: a comparative test. *Applied Psychology: An International Review, 57*, 361–372. http://dx.doi.org/10.1111/j.1464-0597.2007.00318.x.

Judge, T. A., & Kammeyer-Mueller, J. D. (2012). Job attitudes. *Annual Review of Psychology, 63*, 341–367. DOI: 0066-4308/12/0110-0341.

Kittur, A., Nickerson, J. V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., et al. (2013). *The future of crowd work*. Paper presented at the Proceedings of the 2013 conference on Computer Supported Cooperative Work.

Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*, 142–164. http://dx.doi.org/10.1017/iop.2015.13.

Landy, F. J., & Conte, J. M. (2009). *Work in the 21st century: an introduction to industrial and organizational psychology* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Lehdonvirta, V., & Mezier, P. (2013). *Identity and self-organization in unstructured work*. Hertfordshire AL10 9AB, UK: The University of Hertfordshire.

Lemelle, C. J., & Scielzo, S. A. (2012). How you feel about yourself can affect how you feel about your job: a meta-analysis examining the relationship of core self-evaluations and job satisfaction. *Journal of Business Diversity, 12*, 116–133.

Levy, P. E. (2006). *Industrial/organizational psychology: Understanding the workplace* (2nd ed.). Boston: Houghton Mifflin Company.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31.

Loher, B. T., Noe, R. A., Moeller, N. L., & Fitzgerald, M. P. (1985). A meta-analysis of the relation of job characteristics to job satisfaction. *Journal of Applied Psychology, 70*, 280–289. http://dx.doi.org/10.1037/0021-9010.70.2.280.

Maertz, C. P., & Campion, M. L. (2004). Profiles in quitting: Integrating process and content turnover theory. *Academy of Management Journal, 47*, 566–582. http://dx.doi.org/10.2307/20159602.

Marshall, C. C., & Shipman, F. M. (2013). *Experiences surveying the crowd: Reflections on methods, participation, and reliability*. Paper presented at the Proceedings of the 5th Annual ACM Web Science Conference.

Martin, D., Hanrahan, B. V., O'Neill, J., & Gupta, N. (2014). *Being a turker*. Paper presented at the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing, Baltimore, MD.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*, 1–23. http://dx.doi.org/10.3758/s13428-011-0124-6.

Meyer, R. D., & Dalal, R. S. (2009). Situational strength as a means of conceptualizing context. *Industrial and Organizational Psychology, 2*, 99–102. http://dx.doi.org/10.1111/j.1754-9434.2008.01114.x.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: a simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6*, 328–362. http://dx.doi.org/10.1177/1094428103254673.

Nielsen, I. K., Jex, S. M., & Adams, G. A. (2000). Development and validation of scores on a two-dimensional workplace friendship scale. *Educational and Psychological Measurement, 60*, 628–643. http://dx.doi.org/10.1177/00131640021970655.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill Publishing.

O'Neill, T. A., Hambley, L. A., Greidanus, N. S., MacDonnell, R., & Kline, T. J. B. (2009). Predicting teleworker success: an exploration of personality, motivational, situational, and job characteristics. *New Technology, Work, and Employment, 24*, 144–162. http://dx.doi.org/10.1111/j.1468-005X.2009.00225.x.

Paolacci, G., & Chandler, J. (2014). Inside the turk: understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23*, 184–188. http://dx.doi.org/10.1177/0963721414531598.

Paritosh, P., Ipeirotis, P., Cooper, M., & Suri, S. (2011). *The computer is the new sewing machine: Benefits and perils of crowdsourcing*. Paper presented at the 20th International Conference on World Wide Web (WWW), Hyderabad, India.

Ross, J., Irani, L. C., Silberman, M. S., Zaldivar, A., & Tomlinson, B. (2010). *Who are the crowdworkers?: Shifting demographics in Mechanical Turk*. Paper presented at the CHI '10, Atlanta, Georgia, USA.

Ryan, A. M., & Wessel, J. L. (2015). Implications of a changing workforce and workplace for justice perceptions and expectations. *Human Resource Management Review, 25*(2), 162–175. http://dx.doi.org/10.1016/j.hrmr.2015.01.001.

Salehi, N., Irani, L. C., Bernstein, M. S., Alkhatib, A., Ogbe, E., Milland, K., et al. (2015). *We are dynamo: Overcoming stalling and friction in collective action for crowd workers*. Paper presented at the CHI 2015, Seoul, Republic of Korea.

Sauter, S. L., Brightwell, W. S., Colligan, M. J., Hurrell, J. J., Katz, T. M., LeGrande, D. E., et al. (2002). *The changing organization of work and the safety and health of working people*. Cincinnati: National Institute for Occupational Safety and Health.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*, 147–177. http://dx.doi.org/10.1037//1082-989x.7.2.147.

Schmidt, G. (2015). Fifty days an MTurk Worker: the social and motivational context for Amazon Mechanical Turk Workers. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*, 165–171. http://dx.doi.org/10.1017/iop.2015.20.

Schriner, A., & Oerther, D. (2014). No really, (crowd) work is the silver bullet. *Procedia Engineering, 78*, 224–228. http://dx.doi.org/10.1016/j.proeng.2014.07.060.

Schulze, T., Krug, S., & Schader, M. (2012). *Workers' task choice in crowdsourcing and human computation markets*. Paper presented at the 33rd International Conference on Information Systems, Orlando, FL.

Shaw, J. D., Duffy, M. K., Jenkins, G. D., & Gupta, N. (1999). Positive and negative affect, signal sensitivity, and pay satisfaction. *Journal of Management, 25*, 189–206. http://dx.doi.org/10.1177/014920639902500204.

Silberman, M., Irani, L., & Ross, J. (2010). Ethics and tactics of professional crowdwork. *XRDS: Crossroads, The ACM Magazine for Students, 17*(2), 39–43. http://dx.doi.org/10.1145/1869086.1869100.

Snyder, M., & Ickes, W. (1985). Personality and social behaviour. In G. Lindzey, & E. Aronson (Eds.), *Handbook of social psychology: Vol. 2, special fields and applications* (pp. 883–948). New York: Random House.

Staffelbach, M., Sempolinski, P., Hachen, D., Kareem, A., Kijewski-Correa, T., Thain, D., et al. (2014). *Lessons learned from an experiment in crowdsourcing complex citizen engineering tasks with Amazon Mechanical Turk*. Paper presented at the Collective Intelligence Conference, Cambridge, MA.

Staw, B. M., & Ross, J. (1985). Stability in the midst of change: a dispositional approach to job attitudes. *Journal of Applied Psychology, 70*, 469–480. http://dx.doi.org/10.1037/0021-9010.70.3.469.

Summerville, A., & Chartier, C. R. (2012). Pseudo-dyadic "interaction" on Amazon's Mechanical Turk. *Behavior Research Methods, 45*, 116–124. http://dx.doi.org/10.3758/s13428-012-0250-9.

Templer, K. J. (2012). Five-Factor Model of personality and job satisfaction: the importance of agreeableness in a tight and collectivistic Asian society. *Applied Psychology: An International Review, 61*, 114–129. http://dx.doi.org/10.1111/j.1464-0597.2011.00459.x.

Teodoro, R., Ozturk, P., Naaman, M., Mason, W., & Lindqvist, J. (2014). *The motivations and experiences of the on-demand mobile workforce*. Paper presented at the Computer-Supported Cooperative Work and Social Computing, Baltimore, MD, USA.

Tetrick, L. E., Slack, K. J., Da Silva, N., & Sinclair, R. R. (2000). A comparison of the stress–strain process for business owners and nonowners: differences in job demands, emotional exhaustion, satisfaction, and social support. *Journal of Occupational Health Psychology, 5*, 464–476. http://dx.doi.org/10.1037/1076-8998.5.4.464.

Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: path analyses based on meta-analytic findings. *Personnel Psychology, 46*, 259–293. http://dx.doi.org/10.1111/j.1744-6570.1993.tb00874.x.

Thompson, E. R., & Phua, F. T. T. (2012). A brief index of affective job satisfaction. *Group & Organization Management, 37*, 275–307. http://dx.doi.org/10.1177/1059601111434201.

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: a useful supplement to regression analysis. *Journal of Business and Psychology, 26*(1), 1–9. http://dx.doi.org/10.1007/s10869-010-9204-3.

Vakharia, D., & Lease, M. (2015). *Beyond Mechanical Turk: an analysis of paid crowdwork platforms*. Paper presented at the iConference, Newport Beach, CA, USA.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063.

Webster, J., Trevino, L. K., & Ryan, L. (1993). The dimensionality and correlates of flow in human-computer interactions. *Computers in Human Behavior, 9*, 411–426. http://dx.doi.org/10.1016/0747-5632(93)90032-N.

Williams, M. L., McDaniel, M. A., & Nguyen, N. T. (2006). A meta-analysis of the antecedents and consequences of pay level satisfaction. *Journal of Applied Psychology, 91*, 392–413. http://dx.doi.org/10.1037/0021-9010.91.2.392.

Zittrain, J. (2008). Ubiquitous human computing. Philosophical transactions of the royal society of London a: Mathematical. *Physical and Engineering Sciences, 366*(1881), 3813–3821. http://dx.doi.org/10.1098/rsta.2008.0116.

Zyskowski, K., Morris, M. R., Bigham, J. P., Gray, M. L., & Kane, S. K. (2015). *Accessible crowdwork? Understanding the value in and challenge of microtask employment for people with disabilities*. Paper presented at the Computer-Supported Cooperative Work and Social Computing, Vancouver, BC, Canada.