



## Breaking monotony with meaning: Motivation in crowdsourcing markets<sup>☆</sup>



Dana Chandler<sup>a,\*,1</sup>, Adam Kapelner<sup>b,1</sup>

<sup>a</sup> Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, United States

<sup>b</sup> The Wharton School of the University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA 19104, United States

### ARTICLE INFO

#### Article history:

Received 12 May 2010

Received in revised form 3 October 2012

Accepted 2 March 2013

Available online 15 March 2013

#### Keywords:

Natural field experiment

Worker motivation

Crowdsourcing

Online labor markets

### ABSTRACT

We conduct the first natural field experiment to explore the relationship between the “meaningfulness” of a task and worker effort. We employed about 2500 workers from Amazon’s Mechanical Turk (MTurk), an online labor market, to label medical images. Although given an identical task, we experimentally manipulated how the task was framed. Subjects in the *meaningful* treatment were told that they were labeling tumor cells in order to assist medical researchers, subjects in the *zero-context* condition (the control group) were not told the purpose of the task, and, in stark contrast, subjects in the *shredded* treatment were not given context and were additionally told that their work would be discarded. We found that when a task was framed more meaningfully, workers were more likely to participate. We also found that the meaningful treatment increased the quantity of output (with an insignificant change in quality) while the shredded treatment decreased the quality of output (with no change in quantity). We believe these results will generalize to other short-term labor markets. Our study also discusses MTurk as an exciting platform for running natural field experiments in economics.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Economists, philosophers, and social scientists have long recognized that non-pecuniary factors are powerful motivators that influence choice of occupation. For a multidisciplinary literature review on the role of meaning in the workplace, we recommend Rosso et al. (2010). Previous studies in this area have generally been based on ethnographies, observational studies, or laboratory experiments. For instance, Wrzesniewski et al. (1997) used ethnographies to classify work into jobs, careers, or callings. Using an observation study, Preston (1989) demonstrated that workers may accept lower wages in the non-profit sector in order to produce goods with social externalities. Finally, Ariely et al. (2008) showed that labor had to be both recognizable and purposeful to have meaning. In this paper, we limit our discussion to the role of meaning in economics, particularly through the lens of competing differentials. We perform the first *natural field experiment* (Harrison and List, 2004)

<sup>☆</sup> The authors wish to thank Professor Susan Holmes of Stanford University for comments and for allowing us to adapt the DistributeEyes software for our experiment (funded under NIH Grant # R01GM086884-02). They gratefully acknowledge financial support from the George and Obie Schultz Fund and the NSF Graduate Research Fellowship Program. The authors also thank Iwan Barankay, Lawrence Brown, Rob Cohen, Geoff Goodwin, Patrick DeJarnette, John Horton, David Jiménez-Gomez, Emir Kamenica, Abba Krieger, Steven Levitt, Blakeley McShane, Susanne Neckermann, Paul Rozin, Martin Seligman, Jesse Shapiro, Jörg Spenkuch, Jan Stoop, Chad Syverson, Mike Thomas, Abraham Wyner, seminar participants at the University of Chicago, and our reviewers.

\* Corresponding author.

E-mail addresses: [dchandler@mit.edu](mailto:dchandler@mit.edu), [dchandler@uchicago.edu](mailto:dchandler@uchicago.edu) (D. Chandler), [kapelner@wharton.upenn.edu](mailto:kapelner@wharton.upenn.edu) (A. Kapelner).

<sup>1</sup> Both authors contributed equally to this work.

in a real effort task that manipulates levels of meaningfulness. This method overcomes a number of shortcomings of the previous literature, including: interview bias, omitted variable bias, and concerns of external validity beyond the laboratory.

We study whether employers can deliberately alter the perceived “meaningfulness” of a task in order to induce people to do more and higher quality work and thereby work for a lower wage. We chose a task that would appear meaningful for many people if given the right context—helping cancer researchers mark tumor cells in medical images. Subjects in the *meaningful* treatment were told the purpose of their task is to “help researchers identify tumor cells;” subjects in our *zero-context* group were not given any reason for their work and the cells were instead referred to as mere “objects of interest” and laborers in the *shredded* group were given zero context but also explicitly told that their labelings would be discarded upon submission. Hence, the pay structure, task requirements, and working conditions were identical, but we added cues to alter the perceived meaningfulness of the task.

We recruited workers from the United States and India from Amazon’s Mechanical Turk (MTurk), an online labor market where people around the world complete short, “one-off” tasks for pay. The MTurk environment is a spot market for labor characterized by relative anonymity and a lack of strong reputational mechanisms. As a result, it is well-suited for an experiment involving the meaningfulness of a task since the variation we introduce regarding a task’s meaningfulness is less affected by desires to exhibit pro-social behavior or an anticipation of future work (career concerns). We ensured that our task appeared like any other task in the marketplace and was comparable in terms of difficulty, duration, and wage.

Our study is representative of the kinds of natural field experiments for which MTurk is particularly suited. Section 2.2 explores MTurk’s potential as a platform for field experimentation using the framework proposed in Levitt and List (2007, 2009).

We contribute to the literature on compensating wage differentials (Rosen, 1986) and the organizational behavioral literature on the role of meaning in the workplace (Rosso et al., 2010). Within economics, Stern (2004) provides quasi-experimental evidence on compensating differentials within the labor market for scientists by comparing wages for academic and private sector job offers among recent Ph.D. graduates. He finds that “scientists pay to be scientists” and require higher wages in order to accept private sector research jobs because of the reduced intellectual freedom and a reduced ability to interact with the scientific community and receive social recognition. Ariely et al. (2008) use a laboratory experiment with undergraduates to vary the meaningfulness of two separate tasks: (1) assembling Legos and (2) finding 10 instances of consecutive letters from a sheet of random letters. Our experiment augments experiment 1 in Ariely et al. (2008) by testing whether their results extend to the field. Additionally, we introduce a richer measure of task effort, namely *task quality*. Where our experiments are comparable, we find that our results parallel theirs.

We find that the main effects of making our task more meaningful is to induce a higher fraction of workers to complete our task, hereafter dubbed as “induced to work.” In the meaningful treatment, 80.6 percent of people labeled at least one image compared with 76.2 percent in the zero-context and 72.3 percent in the shredded treatments.

After labeling their first image, workers were given the opportunity to label additional images at a declining piece rate. We also measure whether the treatments increase the quantity of images labeled. We classify participants as “high-output” workers if they label five or more images (an amount corresponding to roughly the top tercile of those who label) and we find that workers are approximately 23 percent more likely to be high-output workers in the meaningful group.

We introduce a measure of task quality by telling workers the importance of accurately labeling each cell by clicking as close to the center as possible. We first note that MTurk labor is high quality, with an average of 91 percent of cells found. The meaning treatment had an ambiguous effect, but the shredded condition in both countries lowered the proportion of cells found by about 7 percent.

By measuring both quantity and quality we are able to observe how task effort is apportioned between these two “dimensions of effort.” Do workers work “harder” or “longer” or both? We found an interesting result: the meaningful condition seems to increase quantity without a corresponding increase in quality and the shredded treatment decreases quality without a corresponding decrease in quantity. Investigating whether this pattern generalizes to other domains may be a fruitful future research avenue.

Finally, we calculate participants’ average hourly wage based on how long they spent on the task. We find that subjects in the meaningful group work for \$1.34 per hour, which is 6 cents less per hour than zero context participants and 14 cents less per hour than shredded condition participants.

We expect our findings to generalize to other short-term work environments such as temporary employment or piece-work. In these environments, employers may not consider that non-pecuniary incentives of meaningfulness matter; we argue that these incentives do matter, and to a significant degree.

Section 2 provides background on MTurk and discusses its use as a platform for conducting economic field experiments. Section 3 describes our experimental design. Section 4 presents our results and discussion and Section 5 concludes. Appendix A provides full details on our experimental design and Appendix B is a technical appendix for conducting experiments using the MTurk platform. Both appendices are available online.

## 2. Mechanical turk and its potential for field experimentation

Amazon’s Mechanical Turk (MTurk) is the largest online, task-based labor market and is used by hundreds of thousands of people worldwide. Individuals and companies can post tasks (known as Human Intelligence Tasks, or “HITS”) and have them completed by an on-demand labor force. Typical tasks include image labeling, audio transcription, and basic Internet

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات