Social Media Participation and Performance at Work: A Longitudinal Study

N. Sadat Shami
IBM Corporate Headquarters
1 New Orchard Rd., Armonk, NY 10504
sadat@us.ibm.com

Jeffrey Nichols, Jilin Chen
IBM Research – Almaden
650 Harry Road, San Jose, CA 95120
{jwnichols, jilinc} @ us.ibm.com

ABSTRACT
The use of social media at work is gaining traction, and there is evidence to suggest that various benefits accrue from its use. Yet the relationship between using social media at work and employee performance is not clear. Through a study of 75,747 employees of a large global company over the course of 3 years, we find that some social media usage (number of forum posts, forum post length, and status update length) was positively associated with performance ratings. This study is one of the first to show the relationship among different forms of social media use and employee performance ratings.

Author Keywords
Social media; social software; work; performance**.**

ACM Classification Keywords
H5.3. [Information Interfaces]: Group and Organization Interfaces – *Computer-supported cooperative work*.

INTRODUCTION
Research shows that people's perceptions of the utility of a technology is formed differently when that technology is used in the workplace rather than outside of it [6]. Treem and Leonardi propose an 'affordance approach' showing how social media use at work ranks higher on characteristics of visibility, persistence, editability, and association compared to other forms of communication such as email or instant messaging [7]. This approach may explain why people using the same technology may engage in similar or disparate work practices, and allows researchers to transcend the particularities of any technology or its features, and focus on outcomes.

The affordances of social media may lead to different outcomes based on use and non-use. In this paper, we examine if different forms of social media usage are associated with employee performance, an outcome of interest to many organizations adopting social tools internally. We report the results of a longitudinal study analyzing the relationship between the number of posts of different forms of internal social media and performance ratings. Additionally, we perform topic modeling on social content of top performers to determine if there is a difference in their content compared to others. We conclude by outlining implications for employers and employees.

CHI 2014, April 26 – May 1, 2014, Toronto, Ontario, Canada.

http://dx.doi.org/10.1145/2556288.2557417

RELATED WORK
Not many studies have looked at how social media use in the workplace impacts employee performance. There is however considerable research that demonstrates that employees see value in social media. For example, DiMicco et al. interviewed 17 users of IBM's Beehive social network and found that it was perceived as valuable for making personal and business connections, career advancement, and to support ideas and projects [3]. An annual survey of Microsoft employees from the years 2008 through 2011 found gradual acceptance of social technologies over time, but that a certain degree of skepticism remains [2]. Broadly, these studies have been about the motivation and how social technologies are adopted and used in the workplace. None of these studies include measures of employee performance, and no study has examined employee social media use and performance over time.

Research has investigated employee performance in relation to email and verbal communication. Aral et al. found that in an email network each new piece of information received by a worker resulted in a positive economic gain [1]. Wu et al. analyzed verbal communication and the effect of face-to-face networks on productivity, finding that network cohesion was positively related to productivity [8]. Our work differs from these in that it a) analyzes social media content publicly available within the enterprise instead of other more private communication channels, and b) examines the relationship between different forms of social media use and employee performance.

STUDY
Our study was conducted in a large global company with over 400,000 employees. The company specializes in Information Technology services, software, and hardware. A social platform supporting blogs, forums, and status updates was deployed within the company and used by its employees, but use of the platform was not required nor

was usage of the platform a factor in performance reviews. Because of privacy laws and the sensitivity of performance data, we limited our study to employees working in the United States (US). The company's Human Resources (HR) department anonymized all data, including social media data, before they transferred it to us for analysis.

Our sample consists of 75,747 US employees. Since our goal was to longitudinally track how social media participation relates to performance for the same employees over 3 years, we needed to have data from employees that had been in the company for at least 2.5 years (an employee could receive a performance review based on 6 months of work). Furthermore, it usually takes some time for employees to become acculturated in the large company we studied. So while younger new hires are presumably more comfortable with social media, being new to the company may bias their social media participation compared to other employees that have overcome the acculturation hurdle.

MEASURES

Data for each measure was recorded in 3 separate years: 2010, 2011 and 2012, and used as a control variable in our model. This allows us to control for baseline performance and social media use, as well as year-to-year variability.

Dependent Variable: Performance Rating

Like many companies, the organization studied conducts an annual evaluation of each employee that results in a performance rating. This rating is determined through a deliberative process involving the employee's immediate manager, that manager's manager, and that manager's peer managers. The employee's immediate manager provides an initial ranking and rating for each of his or her employees based on their accomplishments for the year and then each employee is compared to their peers. The final result is a consensus rating agreed upon by all of the managers. Performance ratings have important implications for an employee's career as bonuses and promotion decisions are partially based on them.

In this company, the performance rating is a 4-point numeric scale ranging from 1 for those among the lowest contributors, 2 for average contributors, 3 for above average contributors and 4 for the top contributors.

Independent Variables

The independent variables are all measures of social media usage from the company's internal social media platform. Table 1 shows the number of these posts from our sample over a 3-year period, and their description follows.

Status Updates: A microblogging capability similar to Twitter was available, though these updates are limited to 255 characters. Status updates are visible on an employee's profile page and to those following that employee.

Community Forum Posts: A threaded discussion forum was available for each online community on the platform.

Type of post	2010	2011	2012
Status updates	52,588	74,047	111,721
Community Forum posts	44,730	60,800	83,054
Blog posts	7,256	9,311	11,939

Table 1. Different forms of social posts over time.

Blog Posts: Employees can publish blogs on the platform that are visible to everyone.

We calculate the values for two independent variables for each social media type from the collected data:

Post Count: The total number of posts that an individual made for a given social media type in a given year. This was a repeated measure, and was collected for every employee for 3 years (2010-2012). All 3 types of social media content follow a power law distribution, and were log transformed (base 2, after adding a start value of 1) to control for skew.

Average Post Length: The average character count of posts that an individual made for a given social media type in a given year. This repeated measure was collected for every employee for 3 years (2010-2012).

Dependent variable		**N (%)**
Performance rating [a]	Rating of 1 (low)	6,119 (2.7%)
	Rating of 2 (average)	78,511 (34.5%)
	Rating of 3 (above average)	109,263 (48.1%)
	Rating of 4 (top)	33,348 (14.7%)
Independent variables		
Community forum posts [c]	*M*=0.11, *Mdn*=0, *SD*=0.5, *Min*=0, *Max*=7.97	
Community forum post average length [c]	*M*=249.1, *Mdn*=175, *SD*=241.29, *Min*=2, *Max*=4,063	
Blog posts [c]	*M*=0.02, *Mdn*=0, *SD*=0.2, *Min*=0, *Max*=6.2	
Blog posts average length [c]	*M*=373.8, *Mdn*=239, *SD*=487.88, *Min*=2, *Max*=4059	
Status updates [c]	*M*=0.12, *Mdn*=0, *SD*=0.5, *Min*=0, *Max*=7.8	
Status updates average length [c]	*M*=124.75, *Mdn*=98, *SD*=114.3, *Min*=1, *Max*=255	
Control variables		
Years at company [c]	*M*=16.54, *Mdn*=13, *SD*=10.05, *Min*=2.5, *Max*=57	
Job category [b]	Technical role	115,537 (50.84%)
	Business role	111,128 (48.9%)
	(*Missing*)	576 (0.25%)
Manager? [b]	Non Manager	199,719 (87.9%)
	Manager	24,489 (12.1%)
	(*Missing*)	33 (0.0%)
Age [c]	*M*=48.01, *Mdn*=49, *SD*=9.03, *Min*=21, *Max*=82 (*Missing*=6,713)	

[a.] ordinal variable, [b.] categorical variable, [c.] continuous variable

Table 2. Descriptive statistics. *N* = 227,241

Control Variables

The following control variables are used in our analysis. These variables are measured separately for each employee in each year of analysis.

Years at Company: This variable represents how many years an employee had been at the company at the time of the performance review. For example, an employee that joined in 2009 would have a value of 1 for this variable in the year 2010, and 2 in 2011.

Age: The age of the employee on January 1 of the year in question.

Job Category: The HR department classified 22 job categories. These were divided into Technical role (e.g. software developer, hardware developer etc.) and Business role (e.g. HR, Legal, Finance, Project Executive etc.), which is an officially designated simplified view of the 22 job categories.

Manager: Managers are a special class of employees that may have different patterns of social media usage. This control variable was binary (non-manager and manager).

Table 2 provides descriptive statistics for all variables in their raw form. Using a correlation matrix we found some collinearity, such as between age and years at the company ($r = 0.52$, $p < 0.01$), but not enough to bias results.

RESULTS

We analyzed the data using an ordinal logistic regression model. Generalized Estimating Equations (GEE) is a technique for running statistical analyses for correlated (non-independent) data, such as repeated measures, where the dependent variable is ordinal [5]. When data are collected on multiple occasions on individuals, the responses will not be independent of each other. GEE – an extension of generalized linear models – can account for correlated repeated measurements within individuals, resulting in unbiased regression parameters [4]. Furthermore, GEE models are robust to violations of normality and homogeneity of variance.

GEE with an ordinal logistic model was run with a random factor representing intra-individual correlation per employee and without sensitivity to the entry order of predictors. It was ensured that predictors produced variance inflation factors < 3 to mitigate multicollinearity. Table 3 provides a summary of our analysis. GEE does not provide a goodness of fit measure for ordinal logistic regression.

It is best to interpret GEE in terms of odds ratios. For *one* more forum post, the odds of obtaining a higher level of performance rating increase by 1.003, given all the other variables are held constant ($Exp\ \beta = 1.003$, $p < 0.01$). For *one character* increase in average forum post length, the odds of obtaining a higher performance rating increase by 1.001 ($Exp\ \beta = 1.001$, $p < 0.01$). The average length of internal status updates was significant ($Exp\ \beta = 1.001$, $p < 0.001$), but remaining independent variables were not.

	β (SE)	95% CI for Exp β		
		Lower	Exp β	Upper
Forum Posts	0.002* (0.0009)	1.001	1.002	1.004
Forum Post Length	0.001** (0.0004)	1.000	1.001	1.002
Blog Posts	0.002 (0.0018)	0.998	1.002	1.006
Blog Post Length	0.000 (0.0002)	1.000	1.000	1.001
Status Updates	0.000 (0.0003)	0.999	1.000	1.000
Status Updates Length	0.001*** (0.0002)	1.000	1.001	1.001
Age	-0.036*** (0.0008)	0.963	0.964	0.966
Years At Company	0.015*** (0.0007)	1.014	1.015	1.016
Non Manager	-0.775*** (0.0164)	0.446	0.461	0.476
Job Category				
Technical role	0.527*** (0.104)	1.382	1.693	2.074
Business role	0.376*** (0.1035)	1.189	1.457	1.784

Table 3. Ordinal logistic regression using GEE.
Note: *N*=227,241 observations
*****$p < 0.001$, **$p < 0.01$, *$p < 0.05$**

To explore the underlying mechanisms behind our finding, we analyzed the content of status updates and forum posts created in 2012, the year with the most social media posts. We analyzed the content using Latent Dirichlet Allocation (LDA), using the Stanford Topic Modeling Toolkit. In pre-processing, we removed all posts shorter than 5 words, words that appeared in fewer than 5 posts, words appearing on a 500 stop-word list containing the most common English words, and the most frequent words appearing across all posts. We applied LDA using 5 topics, 10 topics, 20 topics and 30 topics. As the goal of the LDA analysis was to obtain formative insights into the content, we needed to ensure that the generated topics were coherent, meaningful and interpretable, and yet there is no well-established metric to ensure that. Through manually inspecting the resulting topics, we settled on 10 topics for their superior coherence, and assigned each topic a label for description. Our results show that on a majority of topics the distributions are comparable. However an unpaired t-test found that high performers tend to talk more about skills and training, and less on small talk, in comparison to other employees ($p < 0.001$). We report the 10 topics in Table 4, along with their sample words, and the distribution across top performers and everyone else.

DISCUSSION

Our results show that some social media usage, through forums and status updates, is positively associated with performance ratings. We cannot make any claims about causality. Additionally, the effect size is small, which may be ameliorated by selecting a larger unit of analysis. Furthermore, there is much more to performance at work than social media activity. More than anything, our findings

show that increased social participation is not associated with decreased performance, a concern some employers use to justify banning social media use at work.

Topic	Sample topic words (some topics removed to preserve anonymity of research site)	Top performers (rating of 4 and 3)	Everyone else (rating of 2 and 1)
Software development	integration, installation, database, configuration, platform, servers, demo	9.48%	8.18%
Small talk	**folks, probably, pretty, little, person, perhaps, feel, stuff, posts, old**	**15.65%**	**19.78%**
Web authoring	select, pages, field, xml, button, tab, default, folder, format	11.61%	11.49%
Non English content	para, die, und, der, por, das, ich, con, los, ist, den, del, hola, como	2.1%	2.53%
Computer/ mobile	machine, android, firefox, kvm, iphone, vpn, laptop, root, login	11.63%	12.81%
Project report	submission, completed, progress, final, due, phase	8.55%	8.53%
Skills and training	**professionals, training, bright, skilled, efforts, helping**	**5.08%**	**2.5%**
Travel planning	friday, monday, tomorrow, vacation, card, room, august, july	8.55%	8.56%
Leadership and collaboration	leadership, innovation, conference, sessions, workshop, meetings, talk, collaboration, strategy, discuss	14.53%	13.05%
Market analysis	company, technology, products, market, industry, cost, high	12.82%	12.56%

Table 4. Results of LDA topic modeling

In order to speculate why blog posts were not associated with performance, we looked at the aggregated monthly visits to blogs in 2012. Blogs had significantly less visits than forums. They even had less visits than an employee's profile page that displayed her status updates. One of the potential affordances of social media in organizations is visibility [7]. When blogs are not generating enough visits, the benefits that come with visibility such as signaling one's expertise are not realized. On the other hand, the increased visits to forums provide the affordance of visibility to those posting, and may potentially explain the positive relationship between forum posts and performance as a function of expertise and digital reputation. Our study suggests that social media is not monolithic. The affordances of social media for employees are affected by the organizational environment, and may differ based on both the behaviors of others, and the availability of competing forms of social media.

The number of status updates by an employee was not associated with higher performance ratings, while longer status updates were. Perhaps longer status updates tend to contain more business relevant information, whereas shorter updates are more personal in nature. Our topic modeling revealed that top performers differ from others in their use of social media to discuss skills and training, possibly to showcase their skills and signal their desire for continued learning, which may positively affect their performance ratings. Furthermore, status updates largely lack the affordance of persistence [7]. It is often difficult to find old status updates from someone an employee follows.

An implication of our study is that actively posting social media content may be beneficial for employees interested in securing higher performance ratings, but not uniformly. A recommendation for employees is to post content where it would be most visible to others, as well as engage in forms of social media that are more persistent or could be easily retrieved through mechanisms such as search. Meaningful interaction on social media available publicly within the enterprise can signal expertise and allow an employee to showcase her skills to peers and management. This can be amplified further through social sharing.

CONCLUSION

Our study contributes to the literature on social computing through a longitudinal study of employee social media use. An important consequence of our study is the conclusion that social media usage at work does not impede performance. Employers ambivalent about the value of social media at work can take up our findings and encourage adoption of social media among their employees.

REFERENCES

1. Aral, S., Brynjolfsson, E. and Van Alstyne, M. Productivity Effects of Information Diffusion in Networks. In *Proc. ICIS 2007*, (2007).
2. Archambault, A. and Grudin, J. A longitudinal study of facebook, linkedin, and twitter use. In *Proc. CHI 2012*.
3. DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B. and Muller, M. Motivations for social networking at work. In *Proc. CSCW 2008*.
4. Liang, K. Y. and Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 1 (1986), 13-22.
5. Norton, E. C., Bieler, G. S., Ennett, S. T. and Zarkin, G. A. Analysis of prevention program effectiveness with clustered data using generalized estimating equations. *Journal of Consulting and Clinical Psych., 64*, 5 (1996).
6. O'Mahony, S. and Barley, S. R. Do telecommunications technologies affect work and organizations? The state of our knowledge. In B. Staw and R. Sutton (Eds.), *Research in organizational behavior* (Vol. 21, pp. 125-161). Greenwich, CT: JAI Press, 1999.
7. Treem, J. and Leonardi, P. M. Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Communication Yearbook, 36* (2012), 143-189.
8. Wu, L., Waber, B., Aral, S., Brynjolfsson, E. and Pentland, A. Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an IT configuration task. In *Proc. ICIS 2008,* (2008).