

4322 North Quad, 105 S. State St. Ann Arbor, MI 48109-1285

Participation of New Editors After Times of Shock on Wikipedia



Daniel M. Romero School of Information University of Michigan

Co-authored with Ark Zhang, Ruihan Wang, Eric Blohm, Ceren Budak, and Lionel Robert

Collaborative Crowdsourcing

- The Web has enabled large groups of people to collaborate.
- Examples are Wikipedia and GitHub.
- How do people self-organize and adapt to produce high quality output?





Shocks on Collaborative Crowdsourcing



Three Types of Shocks on Wikipedia

Unexpected workforce reduction



ICWSM 2017

Quality Recognition



CSCW 2018

Sudden increase in attention



ICWSM 2019

Sudden Increase in Attention

We identify exogenous shocks of attention on Wikipedia articles corresponding to people:

(i) academics,
(ii) politicians, and
(iii) a random sample of articles
from WikiProject Biography.



We find spikes of attention using Google Trends.

In total, we have 275k articles and 6,662 shocks.

Research Questions

(1) How does the shock affect *collaboration dynamics*?
(2) How does the shock affect *future participation* of newcomers: both directly and through changes in collaboration dynamics?

Collaborative Dynamics

Workload Centralization:

Normalized Gini coefficient of the distribution of number of revisions per editor.



Negative feedback/conflict: Fraction of reverted edits.



Discussion: Fraction of edits to article's talk page and edits to main article.



Impact of Attention Shocks

Attention shocks lead to:

- Increase in activity (more editors and edits).
- Increase in discussion: ratio of talkpage to article edits.
- Increase in reverts.
- Decrease in centralization.

(See paper for details)

Newcomers During Shocks

Attention shocks bring new contributors to the article.

But they they become long term contributors, or quickly stop editing?



Week Number Relative to Shock

Future Participation

We want to measure the future contributions of a group relative to the amount work available.

For an article a in week t, the *future participation* of a group of editors E_t active during week t is:

Num. edits of a by E_t in [t + 1, t + 4]Tota edits in a in [t + 1, t + 4]

Other variants of this definition do not change qualitative results (see paper).

Future Participation

Newcomers



Week Number Relative to Shock

- Newcomers' future participation is generally very low.
- It peaks at around 0.07 at the time of the shocks.
- After the shock, it relaxes back to pre-shock levels of about 0.01.

Future Participation

Newcomers

Incumbents





Week Number Relative to Shock

- It also peaks at the time of shock.
- It remains higher than its pre-shock baseline for several weeks
- This suggests that shocks make articles better at encouraging future participation of incumbents, but not of newcomers.

Research Questions

(1) How does the shock affect collaboration dynamics?(2) How does the shock affect future participation of newcomers: both directly and through changes in collaboration dynamics?



Hypotheses

- 1. Centralization should be **positively** related to future participation.
 - Centralized crowds will be easier to navigate because require less explicit coordination [Kittur and Kraut 2008].
- 2. Reverts of newcomers should be **related** to future participation.
 - Could make newcomers feel unwelcome.
 - Could be interpreted as negative feedback, which encourages productivity [Zhu et al. 2013]
- Discussion should be **positively** related to future participation.
 [Ciampaglia et al. 2013]

Mediation Analysis Set Up

Non - Shock weeks: 2 weeks before the shock Shock weeks: 3 weeks after the shock

Newcomer



centralization = $a + a_1 \times \text{shocks} + \epsilon$ negative feedback to NC = $a + a_2 \times \text{shocks} + \epsilon$ total discussion = $a + a_3 \times \text{shocks} + \epsilon$ discussion per NC = $a + a_4 \times \text{shocks} + \epsilon$

Mediation Analysis Results

Direct Effect	b ₀	1.128***
Indirect Effect		1.045***
Centralization	$a_1 * b_1$	-0.010
Negative feedback to NC	$a_2 * b_2$	0.900***
Total discussion	$a_3 * b_3$	0.135***
Discussion per NC	$a_4 * b_4$	0.000
Total Effect		2.173
Baseline Future Participation	b	8.639

Shocks are associated with an increase in future participation both directly and through:

- Increases in reverts of newcomers (negative feedback).
- Increases in discussion on talkpages (total discussion).

Mediation Analysis Results

Direct Effect	b ₀	1.128***
Indirect Effect		1.045***
Centralization	$a_1 * b_1$	-0.010
Negative feedback to NC	$a_2 * b_2$	0.900***
Total discussion	$a_3 * b_3$	0.135***
Discussion per NC	$a_4 * b_4$	0.000
Total Effect		2.173
Baseline Future Participation	b	8.639

Shocks are associated with an increase in future participation both directly and through:

- Increases in reverts of newcomers (negative feedback).
- Increases in discussion on talkpages (total discussion).

Summary

Attention shocks

- Lead to more activity, reverts, and discussion, and lower centralization.
- Provide an opportunity for articles to recruit new editors.
- Increase future participation both directly and through:
 - Increases in reverts of newcomers (negative feedback).
 - Increases in edits on talkpages (discussion).

Our findings suggest new features that can enable articles to engage newcomers in ways that encourage their sustained participation.

In order to **understand, predict, and manage** social and collaborative systems we have to study them, not only during times of normality, but also **during times of instability**.