

Trust & Wikipedia

Embedding "trust signals" in the desktop reading experience

by mraish + Istigliano / May 2023



Hylaeus adspersus. Image author: Gidip; CC BY-3.0, via Wikimedia Commons.

Key Takeaways

Reader "trust" in Wikipedia encompasses several dimensions, including:

- Trust in the platform as a whole;
- Trust in an individual article;
- Trust in various pieces of information within an article;
- Trust in the author(s) of the article.

Two tested interface cues were associated with more reported trust in Wikipedia—Last Edit Time and Number of Watchers.

Usually, trust is *also* (or exclusively) affected by participants' pre-existing beliefs, experiences, and knowledge of how Wikipedia works.

Last Edit Time

Doeg people

Last edit was on March 17, 2020

Number of Watchers

Salvadoran Civil War

Changes to this article are being watched by 151 editors





Background

project context: "designing for trust and understanding"

The Wikimedia Foundation Web team is exploring the idea of providing Wikipedia readers with more implicit and explicit indications of how the information they consume in Wikipedia articles is generated. By providing contextual information about Wikipedia articles, the Web team hopes to encourage readers to be better informed of the content they are reading and of potential opportunities to participate and contribute.

idea in a nutshell

Include a "trust signal," such as indicating how many editors have contributed to an article, in the desktop reading interface in order to improve both *trust* in Wikipedia and *understanding* of how Wikipedia works.

Karin MacDonald

6 editors have contributed to this article

From <u>Karin MacDonald</u>, one of the source articles used to create study materials.



Stage 2: exposure to "trust signals"

Control (# of languages)	
<i>n=31</i>	文 _A 27 languages ∨
Number of editors	Karin MacDonald
<i>n=32</i>	6 editors have contributed to this article
Number of watchers	Salvadoran Civil War
n=32	Changes to this article are being watched by 151 editors
Last edit time	Doeg people
<i>n=20</i>	Last edit was on March 17, 2020
Last editor name	Doeg people
n=33	Last edit was made by Doug Weller
Open discussions	Salvadoran Civil War
n=29	There are 4 open discussions about this article

Study design



Survey instrument

Participants were sent to a Qualtrics-hosted survey

- Stage 1
 - Demographic and Wikipedia use information;
 - 6 Wikipedia "beliefs and experiences" questions (perceived reliability, trust, accuracy); and
 - 8 T/F questions to demonstrate knowledge of Wikipedia.
- Stage 2
 - Interact with 3 articles—read and answer comprehension questions, some of which are answerable only after having noticed the "trust signal" feature.
- Stage 3
 - Respond to 6 questions about trust in Wikipedia; then
 - When you were reading the articles, did you notice [your trust signal]?
 - Agree or disagree: I trust the information in a Wikipedia article more when I know [the information communicated by my trust signal].



Stage 1: pre-existing stance toward Wikipedia

Question text	measuring
1. When you are in need of information, how confident are you that you can find it on Wikipedia?	Confidence in finding information
2. When you visit Wikipedia, how often do you find the information you are looking for?	Success in finding information
3. Overall, how accurate do you think the information you read on Wikipedia is?	Perceived accuracy of information
4. How confident are you that you can tell when information on Wikipedia is not accurate?	Ability to identify inaccurate information
5. Have you ever found information on Wikipedia that you know was not accurate?	Experience with inaccurate information
6. When you think about the people who add information to Wikipedia, how much do you trust that they are adding accurate information?	Trust in reliability of authors

Stage 2: exposure to "trust signals"

Control (# of languages)	
<i>n=31</i>	文 _A 27 languages ∨
Number of editors	Karin MacDonald
<i>n=32</i>	6 editors have contributed to this article
Number of watchers	Salvadoran Civil War
n=32	Changes to this article are being watched by 151 editors
Last edit time	Doeg people
<i>n=20</i>	Last edit was on March 17, 2020
Last editor name	Doeg people
n=33	Last edit was made by Doug Weller
Open discussions	Salvadoran Civil War
n=29	There are 4 open discussions about this article

Below, you will see part of a Wikipedia article. Please, read it carefully. Once you've read it, you can continue to the next screen where will be asked some questions about

It. = WIKIPEDIA The Free Encyclopedia	Q Search Wikipedia	Creat	te account	
	Doeg people	XA, 6 languages ∨		
Contents [hide]	Last edit was on March 17, 2020	Discussions Article history Edit		
(Тор)	From Wikipedia, the free encyclopedia		On the next screen	
Background "Doag" redirects here. For the Deutsch-Ostafrikanische Gesellschaft (DOAG), see German Ea Frontier The Doeg (also called Dogue, Taux, Tauxenent) ^[1] were a Native American people who lived in Virginia. They spoke an Algonquian language and may have been a branch of the Nanticoke tribe, historically based on the Eastern Shore of Maryland. The Nanticoke considered the Algonquian Lenape as "grandfathers". The Doeg are known for a raid in July 1675 that contributed to colonists' uprising in Bacon's Rebellion. Background [edit] The Doeg (or Dogue) tribe of Virginia were part of the coastal Algonquian language family. They probably spoke Piscataway or a dialect similar to Nanticoke. According to one account, the Doeg had been based in what is now King George County, but about 50 years before the founding of Jamestown (ca. 1557), they spilt into three sections, with groups going to Caroline County and Prince William County, and one remaining In King	ast Africa Company. Deg	Kow much did you know about the Doeg People before reading this article? Nothing at all A title A title A moderate amount		
	The Deeg (or Degue) the of Virginia were part of the coastal Algohquian language family. They probably spoke Piscataway or a dialect similar to Nanticoke. According to one account, the Doeg had been based in what is now King George County, but about 50 years before the founding of Jamestown (ca. 1557), they split into three sections, with groups going to Caroline County and Prince William County, and one remaining in King George. ^{[2],4}	nty, but ions, i King Watercolor by John White depicting an Algonquian village	Alot Monopoint of the topic According to the article, what kind of language did the Doeg People speak? Algorquian	
	When Captain John Smith visited the upper Potomac River in 1608, he noted that the Taux	similar in appearance to villages in Tsenacommacah.	O Muskogean	
	lived there above Aquia Creek, with their capital <i>Tauxenent</i> located on "Doggs Island" (also known as <i>Miompse</i> or <i>May-Umps</i> , now Mason Neck, Virginia.) They gathered fish and also grew corn. Other hamlets were at <i>Pamacocack</i> (later anglicized to "Quantico"), along	Extinct as a tribe Regions with significant populations Vicinia and Mardand	O I don't know	
	Quantico Creek; Yosococomico (now Powells Creek); and Niopsco (Neabsco Creek).	Languages	March 17, 2020	
	Associated with them were other nearby Algonquian peoples — the Moyauns (Piscataway) on the Manuand side, and the Nacotchtank (Anacostan) in what is now the Washington, DC area	Piscataway or Nanticoke (historical)	O January 31, 2017	
	Smith's map also shows a settlement called <i>Tauxsnitania</i> thought to be near present-day	Religion	O June 20, 2022	
	Waterloo in Fauguier County, within the territory of the Siouan-speaking Manahoac tribe.	Native American religion	O I don't remember	
	John Lederer, who visited the Piedmont region of Virginia in 1670, wrote that the entire area had been	Related ethnic groups Nanticoke, Pamunkey, Chickahominy	Survey Completion 100%	
	"formerly possessed by the Tacci, alias Dooi, but the Indians now seated here, are disting	ushed into the several [Siouan] nations of Mahoc.		Continue →

Stage 3: post-exposure trust statements

Question text	measuring
1. Wikipedia is effective in detecting and correcting inaccurate information.	Perceived quality control
2. Knowing who contributes to Wikipedia makes me trust in the information they added.	Perceived reliability
3. It's easy to tell when a Wikipedia article has been modified.	Perceived quality control
4. I know how to tell whether the information provided by Wikipedia is accurate.	Perceived reliability
5. The information provided by Wikipedia is generally accurate.	Perceived reliability
6. If I found inaccurate information on Wikipedia, I would report or correct it.	Perceived quality control
7. If I found inaccurate information in one Wikipedia article, I would trust what I read in other articles less.	Error significance

Analysis questions

To what extent are participants' answers to questions about various aspects of trust in Wikipedia (Stage 3) affected by:

- 1. Their pre-existing knowledge of how Wikipedia works (the Stage 1 T/F questions);
- 2. Their pre-existing "trust stance" toward Wikipedia (Stage 1);
- 3. Their age group and reported Wikipedia access frequency. (Stage 1);
- 4. The specific trust signal that they interacted with in **Stage 2**.



Participants



Participants

- Wikipedia readers recruited from two sources:
 - *Prolific*: 134 participants
 - *Mechanical Turk*: 40 participants
- 62 self-identified female (36%), 102 self-identified male (59%)
- Most participants had completed a bachelor's degree (91; 53%), some had completed high school/secondary (53; 30%), and fewer had completed a master's degree (27; 16%).
- Participant age skews younger—half were between 18 and 25 years old.





Participants

Participants read Wikipedia for a variety of reasons and at a range of frequencies, with roughly half reading Wikipedia on a "weekly" basis.

MR/LS



Problematic Mechanical Turk responses

Most MTurk responses were eventually removed on the basis of responses to the open-ended text questions.

- A large proportion of nonsensical or blank responses
- Generative AI use was suspected on the basis of:
 - Heavily repetitive responses between participants; and
 - A combination of highly accurate and lengthy responses by individuals who had failed a previous closed-ended attention check question.
- <u>This document provides more detail about the problematic responses provided by</u> <u>MTurk participants.</u>



How do you think that information gets added to Wikipedia?

Main themes in responses

1. Volunteers and Users. Well-intentioned volunteers, random people, normal people, "anyone," registered and anonymous users, and volunteer editors.

2. Experts. Subject matter experts, people knowledgeable within the subject they are editing, verified authors, researchers, and professional researchers.

3. Sourcing. Some participants stressed the importance of sources, including verified sources, reputable authors and researchers. Some noted that information can be based on personal knowledge.

4. Passion and Interest. Some participants attribute Wikipedia contributions to people passionate about or interested in a particular subject, including fans enthusiasts, and people with specialized interests.

5. Approval and Verification: Many participants stressed the fact that information needs approval and verification—some mentioned that information is added by people who have been approved to edit or by verified users of Wikipedia.



6. Other groups. Additional categories mentioned include people around the world, teachers, website creators, Google employees, and "Wikipedia staff."

Aggregated findings



WP Knowledge

8 T/F statements about Wikipedia:

- *Most* participants answered *most* questions correctly.
- Largest area of confusion centers on *who is allowed to contribute*.
- Responses used to create a "Wikipedia knowledge" variable (on a 0-8 scale) for each participant.





WP Knowledge





sometimes

usually

always

Pre-exposure: Finding information on Wikipedia





Pre-exposure: Inaccurate information on Wikipedia



Pre-exposure: Accuracy of information on Wikipedia

Post-exposure: Perceived quality control

Post exposure: Perceived reliability

Perceived quality control Perceived reliability

MR/LS

Post-exposure: Error significance

results

All groups noticed their feature at a higher rate than the Control group.

Condition	Did you notice [your feature]?	
	yes	no
Control	16	15
Last Editor Name	25	5
Last Edit Time	15	5
Number of Editors	26	6
Number of Watchers	29	3
Open Discussions	21	8
Total	132	42

Responses to this binary question differ significantly between groups, with all experimental groups responding *yes* more frequently than the Control group. $[\chi^{2(5)} = 15.38, p = .009]$

Did experimental group and/or pre-existing beliefs affect how participants answered questions?

they are adding accurate information?

Factor analysis: Pre-exposure belief questions load onto two new factor variables.* Informatio Inaccurate

Informatio Inaccurate n Access Information and Reliability

Question text	measuring	Factor1	Factor2	Uniqueness
1. When you are in need of information, how confident are you that you can find it on Wikipedia?	1. Confidence in finding information	.84		.29
2. When you visit Wikipedia, how often do you find the information you are looking for?	2. Success in finding information	.72		.48
3. Overall, how accurate do you think the information you read on Wikipedia is?	3. Perceived accuracy of information	.86		.25
4. How confident are you that you can tell when information on Wikipedia is not accurate?	4. Ability to identify inaccurate information		78	.33
5. Have you ever found information on Wikipedia that you know was not accurate?	5. Experience with inaccurate information		.78	.32
6. When you think about the people who add information to Wikipedia, how much do you trust that	6. Trust in reliability of authors	.80		.36

*Exploratory factor analysis (principal component) statistics provided in Appendix.

analysis

Ordered logistic regression—relationship between responses to post-exposure questions and . . .

- Pre-existing knowledge of Wikipedia (the 8 T/F questions)
- Pre-existing attitudes about Wikipedia:
 - Information Access and Reliability beliefs (Factorl)
 - Inaccurate Information beliefs and experiences (Factor2)
- Age group
- Wikipedia access frequency (e.g., daily, weekly, monthly, yearly)
- Experimental condition:
 - Control (no feature)
 - Number of Watchers
 - Last Editor Name
 - Last Edit Time
 - Open Discussions
 - Number of Editors

1. Wikipedia is effective in detecting and correcting inaccurate information.*

- Last Edit Time participants responded with higher agreement.^[coef. = 1.31, p = .035]
- ➡ Agreement with this statement is positively predicted by participant's pre-existing *Information Access and Reliability* attitudes.^[coef. = 1.06, p < .001]
- In the opposite direction, participants who access Wikipedia "monthly", rather than daily, weekly, or yearly, are less likely to agree.^[coef. = -1.43, p = .014]
- Knowing more about how Wikipedia works (higher T/F question scores) is associated with slightly *lower* agreement.^[coef. = -.23, p = .039]

* Ordered logistic regression model: Pseudo R²= .17 (p < .001); LR $\chi^2(14)$ = 67.68. Regression table provided in <u>Appendix</u>.

2. Knowing who contributes to Wikipedia makes me trust in the information they added.*

Agreement with this statement is positively predicted by participant's pre-existing *Information Access and Reliability* attitudes. ^[coef. = .46, p = .008]

* Ordered logistic regression model: Pseudo R²= .08 (p = .002); LR $\chi^2(14)$ = 34.69 Regression table provided in <u>Appendix</u>. Graph represents the relationship between levels of agreement with this statement (ranging from "completely disagree" to "completely agree") and participant scores on the *Information Access and Reliability* factor variable, which combines weighted scores on four pre-exposure Wikipedia stance questions. These two variables are weakly positively correlated (Spearman; *rho* = .22, p = .004).

3. It's easy to tell when a Wikipedia article has been modified.*

- **Last Edit Time** participants responded with significantly higher agreement.^[coef. = 1.41, p = .013]
- ► Agreement with this statement is positively predicted by participant's pre-existing *Information Access and Reliability* attitudes.^[coef. = .45, p = .008]
- At the same time, participants' reported attitudes about and experiences with *Inaccurate Information* negatively predict agreement with this statement.^[coef. = -.49, p = .001]

* Ordered logistic regression model: Pseudo R²= .09 (p < .001); LR $\chi^2(14)$ = 41.59. Regression table provided in <u>Appendix</u>.

4. I know how to tell whether the information provided by Wikipedia is accurate.*

Agreement with this statement was partially predicted by participants' pre-existing beliefs and knowledge about Wikipedia.

- ➡ Agreement is positively predicted by participant's pre-existing *Information Access and Reliability* attitudes.^[coef. = .49, p = .006]
- Experience with and attitudes about *Inaccurate Information* is a significant negative predictor.^{[coef. =} -.91, p < .001]
- Knowing more about how Wikipedia works (higher T/F question scores) is associated with slightly *lower* agreement.^[coef. = -.22, p = .029]

^{*} Ordered logistic regression model: Pseudo R²= .14 (p < .001); LR $\chi^2(14) = 66.78$. Regression table provided in <u>Appendix</u>.

Graph represents the relationship between levels of agreement with this statement (ranging from "completely disagree" to "completely agree") and participant scores on the *Inaccurate Information* factor variable, which combines weighted scores on two pre-exposure questions. These two variables are moderately negatively correlated (Spearman; *rho* = -.41, p < .001).

5. The information provided by Wikipedia is generally accurate.*

- ▲ Agreement with this statement is positively predicted by participant's pre-existing *Information Access and Reliability* attitudes.^[coef. = 1.42, p < .001]
- Conversely, knowing more about how Wikipedia works (higher T/F question scores) is associated with slightly *lower* agreement with this statement.^[coef. = -.28, p = .027]

* Ordered logistic regression model: Pseudo R²= .25 (p < .001); LR $\chi^2(14)$ = 74.51. Regression table provided in <u>Appendix</u>. Graph represents the relationship between levels of agreement with this statement (ranging from "completely disagree" to "completely agree") and participant scores on the *Information Access and Reliability* factor variable, which combines weighted scores on four pre-exposure Wikipedia stance questions. These variables are moderately positively correlated (Spearman; *rho* = .56, p < .001).

I trust the information in a Wikipedia article more when I know [the information provided by the feature I saw].*

Agreement with this statement was only predicted by being in the **Number of Watchers** group—participants who saw this feature responded with significantly higher agreement than other participants.^[coef. = 1.76, p = .001]

* Ordered logistic regression model: Pseudo R²= .13 (p < .001); LR $\chi^2(14)$ = 48.88. Regression table provided in <u>Appendix</u>.

Agreement with two statements was not significantly predicted by any explanatory variables.

Question text

6. If I found inaccurate information on Wikipedia, I would report or correct it.

7. If I found inaccurate information in one Wikipedia article, I would trust what I read in other articles less.

measuring

Perceived quality control

Error significance

Last Edit Time and Number of Watchers emerge

Statement	measuring	Trust feature	Pre-existing
1. Wikipedia is effective in detecting and correcting inaccurate information.	Perceived quality control	LET (+)	IAR beliefs (+) monthly reading freq. (-) WP knowledge (-)
2. Knowing who contributes to Wikipedia makes me trust in the information they added.	Perceived reliability	- - - - - - - - - - - - - - - - - - -	IAR beliefs (+)
3. It's easy to tell when a Wikipedia article has been modified.	Perceived quality control	LET (+)	IAR beliefs (+) II beliefs (-)
4. I know how to tell whether the information provided by Wikipedia is accurate.	Perceived reliability		IAR beliefs (+) II beliefs (-) WP knowledge (-)
5. The information provided by Wikipedia is generally accurate.	Perceived reliability	· · · · · · · · · · · · · ·	IAR beliefs (+) WP knowledge (-)
I trust the information in a Wikipedia article more when I know [the information provided by the feature I saw].	Feature impact	NOW (+)	

Discussion: Two trust signals are associated with different ratings

Reported trust in Wikipedia can be affected by interface cues

Participants seeing Last Edit Time were more likely to agree that they can tell when a Wikipedia article has been modified, and more likely to agree that Wikipedia is effective at detecting and correcting inaccurate information.

Participants seeing **Number of Watchers** were more likely to agree that *knowing how many people are watching an article* improves article trustworthiness.

In most cases, agreement is *also* (or exclusively) affected by participants' pre-existing beliefs, experiences, and knowledge.

Last Edit Time

Doeg people

Last edit was on March 17, 2020

Number of Watchers

Salvadoran Civil War

Changes to this article are being watched by 151 editors

Observations

Last Edit Time and Number of Watchers both include *meaningful numbers*—it's easy for lay readers to interpret this information. On the other hand Last Editor Name may present various name formats (or IP addresses) that may be more difficult to interpret or contextualize, and this feature did not emerge as a predictor of agreement in this study.

Last Edit Time directly highlights the malleability of Wikipedia articles and frames them as dynamic documents—in this study, this feature was associated with greater agreement that it's easy to tell when Wikipedia articles have changed.

Last Edit Time was also associated with agreement that Wikipedia effectively corrects inaccurate information, providing evidence that perceptions of document dynamism and perceptions of document accuracy may be linked.

Number of Watchers (i.e., *Changes to this article are being watched by* [X] *editors*) directly highlights an accuracy mechanism that many readers may be unaware of, and indeed participants who saw this feature directly related this feature to *trusting the information in a Wikipedia article*.

Why did NOW participants agree that Number of Watchers leads to increased trust?*

- → Multiple quality filters
 - I think every editor is a filter so, if it has 151 editors, you have 151 filters that will lead you to know that it's being watched by a lot of experts in the subject.
- → Accuracy
 - If it's watched by various editors it is more likely to be accurate.
- → More eyes = more resistant to vandalism
 - Knowing that an article is being carefully watched makes me think that it's much more immune to being vandalized.
 - Makes me more confident that anyone who may add inaccurate information, whether purposefully or not, will have their edits removed or corrected.
 - The more people watching, the better the chance of maintaining the correct standards in the article.
- → Legitimacy
 - I believe that among all the editors of an article, there is at least one who has a knowledge of the subject, so when there is any addition of incorrect content, there will be immediate correction of this.
 - More people editing an article means more knowledge.
 - The more eyes on it, the more of a consensus there can be.
- → Trust
 - I gain trust after watching the huge number of editors watched.
 - Seeing that more editors are watching solidifies my trust in the accuracy and reliability of the info in the article.

* Quotes are taken from the participants in the NOW condition who report that they "agree" that knowing how many editors are watching changes to an article makes them trust the information in the article more.

Observations

Wikipedia readers arrive with pre-formed beliefs that affect their degree of trust in the platform and in the individual article. For most of this study's post-exposure statements, participants' beliefs served to both increase (in the case of *Information Access and Reliability*) and decrease (in the case of *Inaccurate Information*) their level of agreement.

Knowing more about Wikipedia doesn't necessarily correspond with increased trust—agreement with 3 of the 7 post-exposure trust statements was shown to be *slightly lower* when participants had higher scores on the 8 T/F questions about how Wikipedia works. All of these 3 statements concern Wikipedia's accuracy:

- 1. Wikipedia is effective in detecting and correcting inaccurate information.
- 4. I know how to tell whether the information provided by Wikipedia is accurate.
- 5. The information provided by Wikipedia is generally accurate.

Conclusions

Trust in Wikipedia encompasses multiple dimensions, including trust in an article, trust in a specific piece of information, trust in the platform as a whole, and trust in the author(s). *Some* of these dimensions of trust can be affected by *some* of the interface cues tested in this study, as seen in the responses of participants who saw **Last Edit Time** and **Number of Watchers**.

Trust is also mediated by readers' pre-existing experiences and beliefs—general confidence about the accuracy of Wikipedia and confidence in one's ability to locate information there is sometimes associated with greater trust in the platform. On the other hand, previous experiences with *inaccurate information* on Wikipedia—as well as simply knowing more about how Wikipedia works—can *sometimes* lower the amount of trust that readers have in Wikipedia.

Future directions for research

- What is the relationship between article quality and these trust signals? Article genre? E.g., "last edit was 1 year ago" in a medical article vs a sports article?
- What about variations in the specific information communicated by trust signals? E.g., "last edit was 1 year ago" vs "1 day ago"? How will this interact with article topic or genre?
- How can the 8 T/F questions be improved or expanded to capture more nuance in readers' understandings of how Wikipedia works?
- What is the relationship between perceived trust and perceived usefulness? Does decreased trust lead to decreased usage?

