Converging "Truths"? The Role of Digital Knowledge Platforms for Information Flows across Language Barriers and History

Work in Progress. Please don't quote without permission from the Authors.

Digital platforms are widely believed to entrench polarization and "echo chambers," fostering parallel truths and undermining social cohesion—with potentially serious repercussions for political and economic stability, but also for the propagation of algorithmic biases in AI. In this paper, we provide evidence that the opposite dynamic is also possible: Wikipedia's multilingual infrastructure appears to undo long-standing, language-specific filter bubbles embedded in contested nationalistic narratives of past armed conflicts. Specifically, we show how initially stark differences across communities gradually become harmonized by analyzing how historic battles are portrayed on Wikipedia across more than 50 languages and two decades, focusing on 14 major language editions and over 100 conflict events.

To quantify this process of narrative convergence, we build a novel quarterly panel dataset that links revision histories, language metadata, and structured "battle box" figures (e.g., troop strengths, casualties) to multilingual article text. We quantify narrative distance using cross-lingual Large Language Models (LLMs), we embed articles into a shared semantic space and compute cosine distances from a stable centroid derived from over 50 languages as of 2020. Alongside these semantic measures, we extract numeric conflict data to track factual divergence across language editions.

Despite the battles in our sample having occurred more than 175 years ago, we observe large and systematic differences in early Wikipedia versions: articles in the languages of conflict winners tend to be older, longer, more actively revised, and semantically closer to the multilingual centroid—suggesting narrative centrality or agenda-setting influence. Yet over time, both factual and semantic divergences shrink. Using an event-study design, we find that convergence accelerates sharply following the first appearance of cross-language links and standardized infobox templates. These structural features appear to reduce frictions in comparison and editing, triggering distinctive "spurts" of harmonization across language editions.

Our findings highlight how online platforms can mediate the construction of collective memory in ways that counter the fragmenting dynamics typically attributed to digital media. In contrast to social platforms that amplify division, Wikipedia offers an infrastructure that enables convergence in historically polarized narratives. Methodologically, we demonstrate how multilingual LLMs and automated extraction techniques can be combined to trace the evolution of knowledge, disagreement, and reconciliation across language communities at scale.

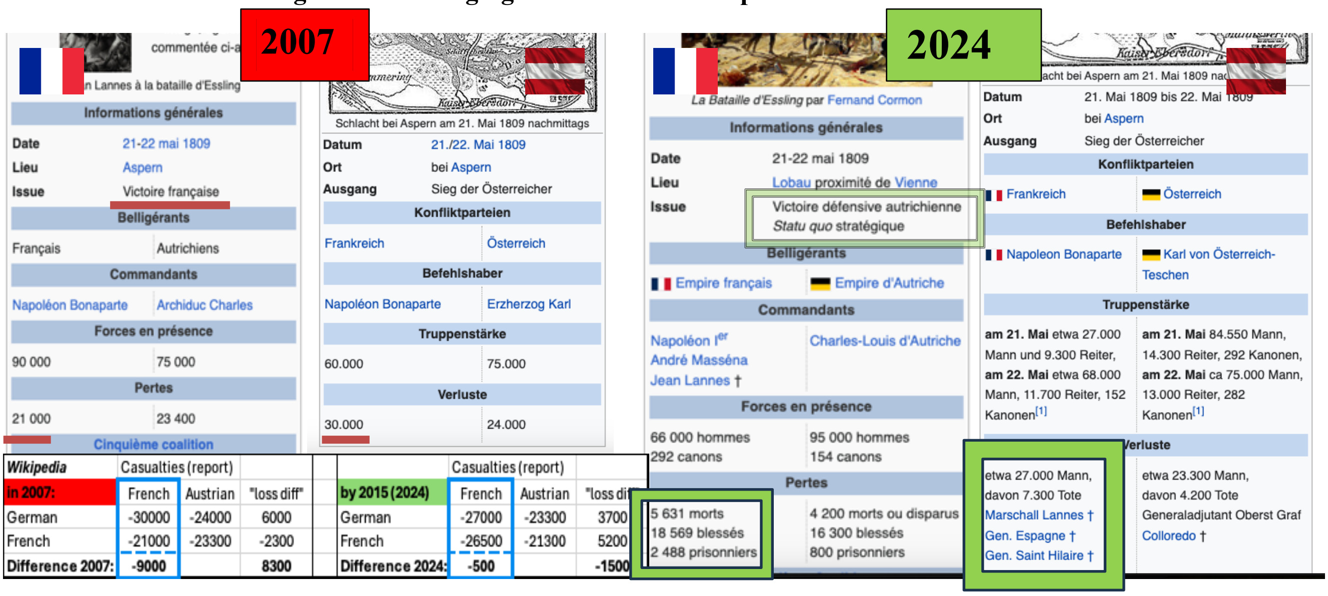
Keywords: Open Knowledge Platforms, AI Translation, Multilingual LLMs, Information Flow, Translation Skills, Misinformation, Polarization, Bias

INTRODUCTION

Discord, polarization and "parallel truths" on digital platforms are widely believed to threaten social cohesion, with potentially serious repercussions on political and economic stability in democratic societies. Because the erosion of mutually agreed 'truth' on social networks could threaten the foundation for societal consensus, this phenomenon has inspired a sizeable body of empirical research about polarization on social media platforms in economics, information systems and other disciplines (Álvarez et al., 2020; Barberá, 2020; Choi et al., 2020; Cinelli et al., 2021; Garimella et al., 2018; Greenstein & Zhu, 2012, Greenstein & Zhu, 2018; Kitchens, Johnson, and Gray, 2020; Terren and Borge-Bravo, 2021). However, the idea of competing and contested narratives is highly familiar in history and memory studies, where cultural memory and historiography as such are construed as fluid and selective processes that are subject to media as well as societal and cultural factors (Halbwachs, 1992; Erll, 2011; Olick, 2011; Assmann, 2011; Garde-Hansen, 2011; Ricoeur, 2004). Conflicts, wars, and pivotal events are understood to be documented from perspectives shaped by nationalistic, cultural, or political frameworks (e.g., Leggewie and Lang, 2011).

Simple examples include differing names for events ("Fall" vs. "Conquest of Constantinople"; "Reconquista" vs. "Fall of Al-Andaluz") and battles with both sides using different field names, casualty numbers or even both sides claiming victory ("Schlacht bei Aspern" vs. "Bataille d'Essling"). For example, older French sources report 21,000 French casualties at the Battle of Aspern-Essling (1809), while Austrian sources state 30,000 French casualties. Similarly, for the Battle of Trafalgar (1805), French accounts cite 26,000 French vs. 18,500 English troops (ratio: 71%), while English sources claim 30,000 French vs. 17,000 English (ratio 54%).

Figure 1: Converging Narratives on Wikipedia: Battle Info Boxes



Notes: Figure 1 contrasts the French and the German Wikipedia "battle info box" summaries for the Battle of Aspern-Essling. The left panel shows the information presented in 2007, the right side corresponds to 2024. The box at the bottom summarizes the respective casualty numbers provided by each side at both points in time. While the divergence in the loss difference is 8300 in 2007, this is reduced to 1500 in 2024 (due to an edit made in 2015). The divergence on French casualties is reduced from 9000 to 500.

We refer to this phenomenon as "*divergence in 'truths',"* because both accounts typically reflect the language-specific tradition of historiography and can be traced back to reputable language-specific encyclopedic sources. Numerous other examples exist of different languages keeping diverging records of historic events of conflict, such as battles and other atrocities (e.g., Leggewie and Lang, 2011). but this divergence has, as of yet, not been studied quantitatively and at large scale. Yet, such biases, deeply embedded in historical accounts foster misunderstanding and could hinder both reconciliation and even economic development. Moreover, such discrepancies in online accounts are especially concerning given that large-scale AI models are increasingly trained on these texts. Biases embedded in historical narratives may thus propagate into algorithmic bias, with potentially harmful consequences for individuals (Fu et al., 2020, 2021). Understanding and quantifying these divergences is therefore critical across domains.

For the example of Aspern-Essling, it is noteworthy that Austrian history refers to it as "Schlacht bei Aspern" (e.g.: 'Lion of Aspern' in Vienna), while French history refers to it as "Bataille d'Essling" (e.g., 'Arc de Triomphe' in Paris). Initially, neither Wikipedia page mentioned the other name, but as of 2009, both do. Further, until 2007, "Bataille d'Essling" stated a French victory, while "Schlacht bei Aspern" stated an Austrian victory. Since 2007, "Bataille d'Essling" describes it as an "Austrian victory without strategic consequences" and the divergence in casualties has shrunk from 9000 to 500 (now: "Bataille d'Essling": 26,500, "Schlacht bei Aspern": 27,000).

While the existing literature has focused on the potential of digital platforms to amplify discord and increase conflict, the present research posits that connecting and juxtaposing conflicting historic narratives can foster a process of harmonization and contribute to a resolution and convergence of long-standing conflicting narratives. We thus depart from the existing "one-way paradigm" of amplified conflict to analyze whether digital platforms also offer novel opportunities to enable a mechanism of reconciliation. However, we stress that the hypothesized mechanism is fundamentally different from solutions suggested in the existing literature on echo chambers (e.g. Terren and Borge-Bravo, 2021).

To study how digital platforms may aid informational flow and reconciliation of "truths", we systematically analyze how historical conflicts are represented across languages on Wikipedia. Wikipedia offers an ideal setting: every edit is publicly recorded (Halatchliyski et al., 2016; Kummer et al., 2021; Hinnosaar et al., 2022, 2023), and its structure enables both language-specific divergence and cross-lingual comparison. This creates a context of clearly connected (via topic and language links) yet unambiguously separated (across languages) narratives—perfectly suited for our analysis. In this sense, Wikipedia provides a unique laboratory to study two key phenomena: (i) the persistence and propagation of language-specific narratives—what we term "historic echo chambers"—and (ii) the platform's potential to foster narrative convergence, or "converging truths."

We tackle these questions by compiling a novel dataset that traces the evolution of Wikipedia articles on over 100 notable battles, drawing on multilingual revision histories from more than 50 languages over the past two decades. This unique setting allows us to examine the formation of historical narratives at scale, taking advantage of the fact that each language edition is edited independently, while sharing a common structure and increasingly standardized conventions. For structured analysis, we extract two types of information: narrative text and factual metadata from semi-structured "battle boxes," which summarize figures such as troop strengths and casualties.

We introduce two complementary measures of cross-lingual disagreement. First, we quantify factual divergence by comparing reported numbers of casualties and troop sizes across language editions, using a custom parser for infobox templates. Second, we compute semantic divergence by embedding article texts in a shared multilingual space using Large Language Models (LLMs), measuring each version's cosine distance from a stable 2020-based centroid derived from over 50 languages. Together, these measures capture both concrete factual contradictions and broader narrative distance. We augment these with metadata on article quality (e.g., age, revision length, and links) and the role of each language community in the conflict (winner, loser, or uninvolved).

We find strong evidence of convergence in how historic battles are portrayed across languages. Despite all battles in our sample occurring over 150 years ago, early Wikipedia versions exhibit considerable disagreement—both in reported facts and in narrative framing. Specifically, we observe large factual discrepancies and greater semantic distances between language editions in earlier versions. Articles in the languages of conflict winners appear earlier, are longer, more frequently edited, and more semantically aligned with the multilingual centroid—suggesting narrative centrality or influence. Yet over time, both factual and semantic divergences decline significantly, pointing to a process of narrative harmonization or "convergence in truths." To explore the drivers of this convergence, we use an event-study design centered on two structural platform features: (i) the introduction of interlanguage links connecting parallel articles across editions, and (ii) the adoption of standardized infobox templates. Both events are associated with marked and lasting reductions in cross-lingual divergence, particularly for initially inconsistent accounts.

Taken together, our findings highlight Wikipedia's capacity to mediate and align contested historical narratives across cultures. Language editions grow more aligned over time in both content and tone, underscoring the integrative role of Wikipedia's multilingual infrastructure. In contrast to the widely studied dynamics of polarization and fragmentation on social media, we document a platform-induced convergence that appears to mitigate—rather than amplify—historical disagreement. Methodologically, our study showcases how multilingual LLMs and automated extraction techniques can be combined to quantify cross-lingual disagreement, trace the evolution of collective memory, and examine platform-induced narrative alignment at scale.

Our framework provides a scalable methodology to quantify convergence in contested narratives and opens new avenues for comparative research across domains. Future extensions will explore less conflictual topics—such as scientific topics or sports events—to enable causal designs like difference-in-differences. We also plan to broaden the scope to additional languages and a more diverse set of historical and contemporary events, further refining our understanding of how digital platforms shape collective memory across contexts.

LITERATURE REVIEW

The potential for bias in online platforms is well-documented. Prior literature highlights in-group bias across languages (Álvarez et al., 2020) and political slants away from neutrality in both expert-curated Encyclopaedia Britannica and crowd-sourced Wikipedia (Greenstein & Zhu, 2012, 2018). We conceptualize historic biases and diverging accounts of historical events between languages and nations as a "historic echo chamber effect," where cultural perspectives and preferences amplify differences in narratives.

The echo chamber effect has been extensively explored in social media, revealing mechanisms such as algorithmic filtering, self-selection, and homophily that reinforce pre-existing beliefs (Cinelli et al., 2021). Studies further show how social media induces partisan shifts, amplifies rumors, and reduces bipartisanship (Kitchens et al., 2020; Choi et al., 2020; Barberá, 2020; Garimella et al., 2018). These dynamics contribute to polarization and challenge democratic discourse, while mitigation strategies, such as diversifying information sources, have been explored (Terren & Borge-Bravo, 2021). These findings resonate with cornerstone insights from the disciplines of history and memory studies who highlight that memory, the recollection of facts, and historiography are fluid and selective processes that are subject to societal and cultural factors (Halbwachs, 1992; Erll, 2011; Olick, 2011.), which interact in complicated ways with the medium that creates the record. (Assmann, 2011; Garde-Hansen, 2011; Ricoeur, 2004).

Theories of truth—coherence, consensus, and pragmatic—offer frameworks for analyzing information convergence. Coherence theory emphasizes logical consistency within interconnected beliefs (Rescher, 1973; Kirkham, 2001), while consensus theory defines truth through rational agreement in a community (Habermas, 1984; Rescher, 1993). Pragmatic theory ties truth to practical consequences, viewing it as utility-driven and evolving with new information (James, 1907; Peirce, 1878). Together, these theories inform how data from multiple sources can merge into a unified narrative (Blanshard, 1939; Jenkins, 2006). The present research posits that connecting and juxtaposing conflicting historic narratives can foster a process of harmonization and contribute to a resolution of conflict. We argue that digital platforms offer a novel opportunity for bringing this mechanism to fruition. However, we stress that the hypothesized mechanism is fundamentally different from solutions suggested in the existing literature on echo chambers (e.g. Terren and Borge-

Bravo, 2021). Wikipedia offers a unique laboratory to study and quantify both phenomena: (i) the existence, transmission, and propagation of language-specific narratives as an instance of "historic echo chambers" and (ii) the extent to which digital knowledge platforms can (or fail to) induce a process of convergence in narratives or "converging truths."

The value of understanding this channel cannot be overemphasized, as historical grievances and differing interpretations of history have been sources of tension and conflict between ethnic groups and nations (see e.g., Galtung, 1996). In contrast, a shared vision of history can foster cooperation and diplomacy and promote stability and peace. (e.g., Larson & Shevchenko, 2010, Montville, 1990). This perspective has also been at the heart of famous efforts including the Gacaca courts in Ruanda (Clark, 2010; Minow, 1998) or truth and reconciliation commissions (Doxtader, 2009; Tutu, 1999; Wilson, 2001). Related research found that providing students with a cohesive understanding of their country's past promoted national unity, social cohesion, and a sense of identity (e.g., Noddings and Brooks, 2016), while shared historical narratives facilitate cultural exchange and tourism between nations. (e.g., Timothy & Boyd, 2006). We take this as overwhelming evidence for the value of promoting a shared narrative of historical accounts.

Wikipedia is ideal to study whether digital knowledge platforms can induce convergence in narratives, because Wikipedia constructs knowledge collaboratively, with consensus emerging through, coherence, and practical relevance (Kittur & Kraut, 2008; Jemielniak, 2014; Reagle, 2010). Talk pages foster dialogue, aligning with consensus theory (Habermas, 1984), while coherence is ensured by integrating new information with existing content (Forte & Bruckman, 2008). Pragmatic utility guides contributions to prioritize relevance (Fallis, 2008). Mechanisms like Creative Commons licenses (Lih, 2009), verifiability policies (Magnus, 2009), and collective editing on talk pages balance differing viewpoints to create unified, unbiased narratives (Forte et al., 2012; Reagle, 2010).

Moreover, Wikipedia's multilingual structure uniquely reflects diverse cultural perspectives, with independently operated language editions often differing in scope and historical interpretations (Hecht & Gergle, 2010). Tools like interlanguage links, multilingual editors, and translation aids, supported by Wikidata, help bridge gaps, streamline updates, and enhance synchronization while highlighting cultural differences (Adafre & de Rijke, 2006; Miquel-Ribé & Laniado, 2018; Vrandecic, 2012; Samoilenko et al., 2017). On the other hand, research has shown that the process is subject to challenges, including resource disparities, cultural biases, and misinformation. Smaller editions lag in updates (Lewoniewski et al., 2017), while edits may reflect national pride or differing sources (Pfeil et al., 2006). Wikipedia's inclusivity also invites unreliable edits, requiring strong editorial practices to ensure verifiability and consistency (Ford et al., 2013; Luyt & Tan, 2010). Language links aid synchronization and information transfer, but full convergence remains limited by resource and contextual differences (Adar et al., 2009; Callahan & Herring, 2011; Liu et al., 2018). To summarize, our work builds on and contributes to four connected streams of literature:

- **Digital Platforms, Echo Chambers and Bias**: Studies highlight the prevalence of biases in digital platforms like social media platforms, and Wikipedia, particularly in early article versions, with later revisions often reflecting more balanced perspectives (Greenstein & Zhu, 2018; Kim et al., 2016). However, systemic biases remain, shaped by editorial practices and cultural perspectives.
- **Historiography and Narrative Bias**: Historical accounts, especially those documenting conflicts, frequently reflect the perspectives of victors or dominant cultural groups (Bloch, 1954; Carr, 1961; Tosh, 2015). This reinforces inequalities in representation and complicates reconciliation efforts (Halbwachs, 1992; Hunt, 1989; Minow, 1998; Tutu, 1999).
- **Peace-building and Shared Histories**: Research which underscores the importance of establishing shared historical narratives in fostering reconciliation and empathy (Bekerman & Zembylas, 2011; Galtung, 1996). Digital platforms represent a novel avenue for promoting such shared understandings, and thus contribute to peace building and collaboration between language groups and societies (Clark, 2010; Minow, 1998; Tutu, 1999; Zehr, 2002).
- **Recent progress in Large Language Modelling** has fueled AI-supported translation and Natural Language Processing methods. Major progress in NLP was achieved by several algorithms, including word2vec (Mikolov et al., 2013), BERT (Devlin et al., 2018), and, most recently, Multilingual Embeddings (Cohere, 2024). Cohere, founded by former Google Brain employees in 2019, focuses on LLMs enabling access to researchers through its nonprofit research lab and open-source community, Cohere for AI.

We identify three key stylized facts that will guide our analysis. First, conflicting parties report events differently, often highlighting their own achievements while downplaying those of their adversaries. Second, the winning side tends to provide more detailed accounts, leveraging its position to shape the narrative—though this is less evident in cases involving crimes against humanity. Finally, discrepancies in accounts, including numerical data such as casualty figures, are common, reflecting the subjective nature of historical reporting. These discrepancies often stem from biases, unequal access to information, or differing political and cultural agendas. For example, casualty figures may vary depending on whether reported by the victor, aiming to downplay losses, or the defeated, highlighting the damage inflicted.

Before moving to the general strategy of the analysis, we stress that the example from the Battle of Aspern-Essling is not unique. A similar pattern (though under opposite signs) emerges for the Battle of Wagram which ensued only 6 weeks later. Other prominent examples are the “Fall”/”Conquest” of Constantinople and of Al Andaluz/Spain, or the Battle of Trafalgar and many more. Our main goal is to analyze these patterns systematically, while our second major goal is to shed light on the relative positioning of “winning”, “losing” and uninvolved languages.

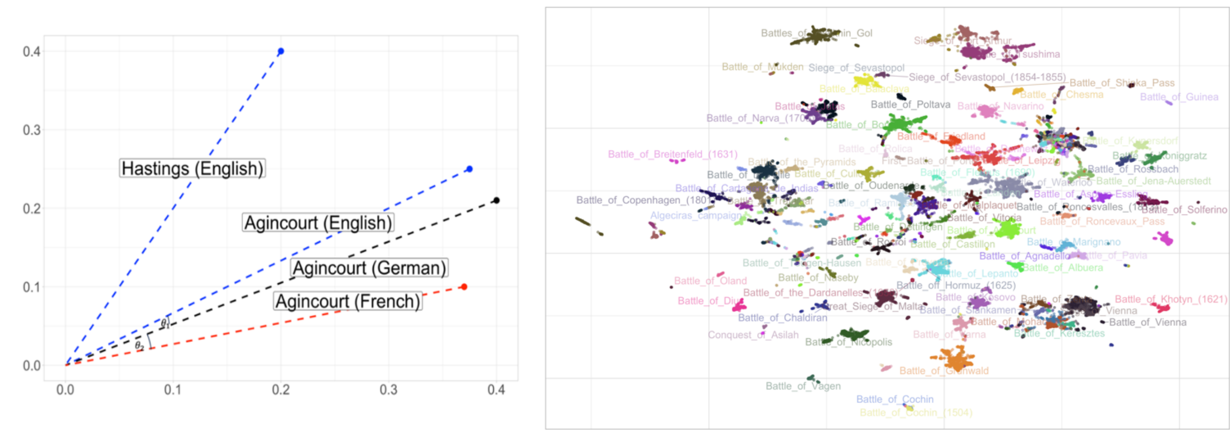
RESEARCH STRATEGY

Main Questions: Our analysis is shaped by three guiding questions, emerging from the examples presented above. First, do biases in reporting grow or diminish over time? Understanding this dynamic can reveal whether historical narratives tend to stabilize or diverge over time. Second, does interlinking on platforms foster a "convergence of truths," where conflicting perspectives align through shared access and connectivity? Third, do these biases, deeply rooted in cultural and historical contexts, persist and transmit through centuries, shaping how societies remember and interpret the past? In addition, we ask what the role of "neutral" languages is in bridging these divides and mediating between conflicting narratives?

Multi-Lingual LLM Embeddings and Cosine Similarity: The core idea is to analyze how the semantic differences in historical narratives across languages evolves over time. To achieve this goal, we leverage multilingual large language model (LLM) embeddings. The core idea is to embed all language versions of an article into one multilingual semantic space and to quantify the alignment or divergence of narratives. The quantification is achieved by evaluating their cosine similarity, a measure of how closely the embedding vectors for two texts (i.e., t_1 and t_2) are aligned in a high-dimensional space. Cosine similarity ranges from 1 (exactly identical) to -1 (exactly opposite) and is equal to one minus the cosine distance.

$$S_{cos}(t_1, t_2) = \frac{\sum_{i=1}^{n} t_{1i} t_{2i}}{\sqrt{\sum_{i=1}^{n} {t_{1i}}^2} \sqrt{\sum_{i=1}^{n} {t_{2i}}^2}}$$

Figure 2: Illustration of Cosine Similarity and Mapping of Cross-Lingual Similarity Scores



This is illustrated in the left panel of Figure 2 with a visual representation based on two historical battles: the Battle of Hastings and the Battle of Agincourt. The figure contrasts semantic cross-language similarity in narratives with same-language semantic distance between two topics (e.g. different battles). In this illustration, the English account of the Battle of *Hastings* shows significant semantic distance to all narratives about the Battle of *Agincourt*, as reflected in its isolated position in the figure. Among the more closely aligned multi-lingual accounts of the Battle of Agincourt the English and German versions show closer alignment, while the French version diverges more significantly. This illustrates how the proposed methodology will allow to quantify how historical events are framed differently across languages, shaped by cultural or linguistic context. For example, the French narrative of Agincourt may emphasize aspects that differ from those in the German or English accounts, reflecting the perspectives of each nation's historiography.

Initial Cross-Language Comparison via LLMs (2023 version): The right panel of Figure 2 presents the results of embedding historical battles into multi-lingual large language model (LLM) by Cohere and visualizing the relationships using dimensionality reduction techniques. The analysis explores how historical narratives differ across languages, with each color representing a specific battle (e.g., Battle of Agincourt, Battle of Vienna) and each point representing the text of a corresponding Wikipedia article in a different language (German, French, English, Swahili, etc.) as embedded by Cohere's multilingual LLM. The spatial distribution shows that articles on the same battle (but from different languages) are strongly clustered, which highlights that the embedding vectors encode meaningful semantic differences. Tightly grouped battles suggest a shared or convergent narrative across linguistic boundaries, while dispersed points highlight divergence in how different languages frame the same event. Some battles, like Battle of Vienna, appear closely aligned across languages, reflecting greater semantic agreement. Others, like Battle of Tsushima or the Siege of Port Arthur, are more dispersed, indicating greater variation in how the event is described in different languages.

Our analysis starts with the cosine similarity of winners (*w*) and losers (*l*) to the topic level centroid of all languages (D_{Awlt}), which we track over time. We estimate linear and flexible time-trends, to evaluate whether divergences grow or diminish over time. We plan to contrast the battle specific similarities with analogous similarity measures for articles about neutral baseline topics, such as planets and chemical elements. A simple specification to estimate group specific linear trends at the topic level *A* is given by:

$$D_{Awlt} = \beta_0 + \beta_1 Battle_{Awl} + \beta_2 t + \beta_3 Battle_{Awl} \times t + \varepsilon_{At} \quad (1)$$

D_{Awlt} is the measure of narrative distance between the article *A* in language of the winner *w* and the loser *l*. $Battle_{Awl}$ takes the value 1 if the article is about a battle and if it was fought between *w* and *l*, and t is a simple numeric variable representing time. Suitable non-parametric alternatives will be estimated using time as a categorial variable and estimating non-parametric methods including splines or kernel smoothing.

The analysis of D_{Awlt} will be underpinned by an analysis of the alternative distance measures that are given by the differences in article properties (e.g. length) and casualty reports (compared against numbers in the info boxes of the control group). In addition to the control groups based on planets and chemical elements, we can also use battle pages for the same languages but about other battles that did not involve *w* and *l* as control groups.

In step 2, we can leverage the availability of languages of uninvolved ("neutral") parties. Using the centroid (i.e. average) of the measure obtained for these languages, we repeat the analysis but contrast the winner's distance to the centroid, D_{AwNt}, to the loser's distance to the centroid, D_{AlNt}. Doing so allows us to analyze initial distances and convergence between each party and the position of the average uninvolved party. Moreover, the contrast of the convergence between neutral and winner versus convergence between neutral and loser will shed additional light on the convergence results (or absence thereof) in D_{Awlt}. A third aggregate measure is given by the circumference of the triangle $D_{Awlt} + D_{AwNt} + D_{AlNt}$, which can be conveniently tracked over time.

Using multi-lingual LLM embeddings, like those from Cohere, combined with Wikipedia's revision history to track changes at a monthly level, represents a novel method for studying narrative convergence and divergence across languages over time. By moving beyond surface-level differences in wording to capture deeper shifts in meaning and the evolving convergence of "truths," this method provides unparalleled insights into how digital platforms can foster convergence in long-standing, diverging historical narratives. Unlike the familiar discord and divergence often seen on social media, this approach explores how such platforms can constructively reshape collective understandings of the past and how language, culture, and economic factors influence this process.

Analysis of baseline convergence and platform interventions: We further analyze knowledge conflicts across languages as well as the efficacy of platform strategies such as cross-language linking or AI-driven translation tools in fostering narrative convergence. As before, we first focus on the analysis of historic accounts of the same battle across different languages and whether initial discrepancies in the different accounts on Wikipedia diminish over time (or do not) because of, among other things, exposure to the other language versions about the same battle.

We will leverage three major design features that were introduced by the platform since 2002 and test how (i) interlinking, (ii) user interface standardization (battle info boxes), and (iii) AI-driven tools for translation affect biases and foster alignment. For all three interventions, this part will deploy event study techniques, and thus a short description of the interventions is in order.

The first intervention, "battle info boxes," are standardized templates summarizing battles in the top right corner of articles (e.g., "Victoire francaise" vs. "Austrian victory" or casualty discrepancies). These templates make stark discrepancies more visible, potentially triggering convergence. Battle info boxes are user-added templates, enabling two event studies: (1) Template Availability: When the template becomes available across a language, allowing estimation of an ITT (Intention to Treat) effect. (2) Template Activation: A battle info box is added to a specific article, allowing ATT (Average Treatment Effect on the Treated) estimation. Identification arises as different language versions (e.g., Dutch or Swedish) activate the battle info box at different times, enabling comparisons across language-battle observations. In addition, uncontested content, such as on geometric shapes or chemical elements, serves as alternative comparison group.

The second intervention, language links, connects articles to their counterparts in other languages, making cross-language versions one click away. Like battle info boxes, language links make discrepancies more visible, triggering convergence. This intervention also supports ITT estimation (when the feature is generally available) and ATT estimation (when links are added). As before, control observations are provided by other language versions and by control topics. The third intervention, AI translations, is rolled out universally across the platform, making treatment exogenous. However, unlike the first two interventions, other languages cannot serve as controls. Instead, we use uncontested content, such as geometric shapes or celestial bodies, as a comparison group to control for baseline differences between languages. This comparison group is also applicable to the first two interventions.

Estimation strategy: Each intervention will be studied independently, but the estimation methodology is consistent across them. We define "TreatAct" as an indicator variable equal to 1 when a feature (e.g., battle info box) is added to an article or article pair, signifying active treatment. While features can be removed, they are typically improved over time. We hypothesize that treatment accelerates convergence, with the null hypothesis being no effect on narrative convergence. To test this, we estimate a DiD regression using article-level outcomes (e.g., article length, cosine similarity, or – for battles – info-box numbers). For simplicity, we focus on battle analysis using planets as a control group and language links as treatment. The methodology generalizes to other topics like historical figures. Our sample includes all battle-language and planet-language observations, forming a topic-language-month panel where an article A in language L is the observational unit. For article length, we construct a balanced monthly panel, coding non-existent articles as zero. For other outcomes, the panel is balanced for existing articles (cosine similarity) or reported casualties (casualty differentials). The simplest two-way fixed effects:

$$\Delta D_{Awlt} = \beta_0 + \beta_1 TreatActive_{Awlt} \times Battle_{Awl} + \delta_A + \gamma_t + \varepsilon_{ALt} \quad (2)$$

This regression can be expanded to a full event study design with period-battle interactions and time normalized to 0 at the time when treatment between *w* and *l* is activated. D_{Awlt} is the measure of narrative distance between the winner's (*w*) and the loser's (*l*) version of article *A* and the centroid of the neutral languages in month *t*. For the event-study analysis we focus on the change in narrative divergence by taking the first difference to the month before. $TreatActive_{ALt}$ equals 1 if and only if treatment is active for the given article *A* at time t (e.g. if an info box is present or a language link is in place) and, otherwise, equals 0. $Battle_{Awl}$ is an indicator variable taking the value 1 if the article is a battle-observation involving languages *w* and *l* as conflicting parties, and 0 if it is from the control group. Article δ_A and weekly fixed effects γ_t are also included. As before, we can expand to compare the winner's and loser's distance to the neutral language centroid (N_A) for article *A* (D_{AwN_At} vs D_{AlN_At}) and include article-language level fixed effects α_{AL} for the winner and the loser respectively (not shown here).

These analyses contribute to the literature on information systems, because each intervention uses a different channel to make conflicting information more salient. Language links improve access to other versions, battle info boxes highlight key contradictions (e.g., in casualties), and AI translations expand the pool of users accessing foreign-language information. Comparing how the effects differ for each channel offers insights into their relative importance in exposing conflicting narratives.

Specification issues: A key challenge in our analysis is the potential endogeneity of language links and battle info boxes, as they are placed by users. First, reverse causality may occur if a link is added due to prior convergence. To address this, we use changes in narrative divergence as the dependent variable and run regressions with increasingly small time windows around the event, ensuring convergence is not misattributed. Pre-trend analyzes further confirm that convergence does not precede the intervention. Edits co-occurring with the placement of links will be studied in detail to understand the role of individual editors versus the crowd.

Second, the user who places a link may also edit articles, systematically reducing divergence. To account for this, we: (1) exclude convergence attributable to the user who placed the link, (2) exclude edits occurring within a week or month of the intervention and focus on changes in convergence speed beyond this window, and (3) conduct "honest DiD" analyzes of pre-trends (Rambachan and Roth, 2019, 2023).

Lastly, battle info boxes and language links are introduced at varying times, creating heterogeneous treatment timing. Standard Two-Way Fixed Effects (TWFE) estimators may produce biased results due to improper weighting of treatment effects, particularly when earlier-treated groups act as controls. To address this, we apply the Callaway and Sant'Anna (2021) methodology, which adjusts for heterogeneous treatment effects and avoids bias from incorrect control group weighting.

DATA AVAILABILITY AND PROOF OF CONCEPT (PILOT STUDY)

Wikipedia provides all the necessary information in the revision history and, more usefully for scraping, in an API which is publicly documented. The API allows to pull the raw text and metadata on every revision through a unique language specific revision ID. From the raw text it is also possible to obtain historical information on the presence of language links. Articles about the same topic (e.g. battle) are connected through a unique wikidata ID, connecting all language versions. (e.g. https://www.wikidata.org/wiki/Q541626 for the battle of Aspern-Essling). Page views and User profiles are publicly provided by wikimedia's x-tools (https://xtools.wmcloud.org/) which also ensures public data availability through an API.

Since data collection is at the core of this proposal we ran two pilot data collections, with data on 46 and 93 battles in up to 14 languages. We focus on pre-WWI battles to avoid complexities of modern warfare (month-long battles, complex equipment) and to minimize potential survivorship bias. A focus on human-related casualties ensured a manageable scope, consistent data and kept the dataset aligned with the goals of the pilot. Our larger sample includes over 200,000 Wikipedia articles from 85 languages on 93 historic conflicts spanning the 14th to 20th centuries, edited by 56,000 editors. Among them, 1,721 edited in more than one language, with 76% contributing to English Wikipedia, underscoring its role as an informational aggregation language (Kim et al., 2016). German Wikipedia was second, with 479 multilingual editors (28%). For each conflict, we identified the involved countries and their languages, finding 33 distinct nations, most prominently France. While English Wikipedia covered all 93 conflicts, many other languages had far fewer articles, highlighting the significant barrier language poses to information flow. Future expansions will address underrepresented regions to improve coverage.

To prepare articles for embedding, the articles are cleaned and batched before being run against the Cohere v1/embed-jobs API endpoint (Team et al., 2023). This process enables clustering, distance, and topic analysis to assess text (dis)similarity. Cosine similarity is applied to derive topic- and language-level averages, identifying the most divergent topics and languages. We also plan to use cluster analysis techniques like t-SNE and Uniform Manifold Approximation and Projection (umap) for visualizing variation (Maaten & Hinton, 2008) and apply language fixed effects to account for differences in information density.

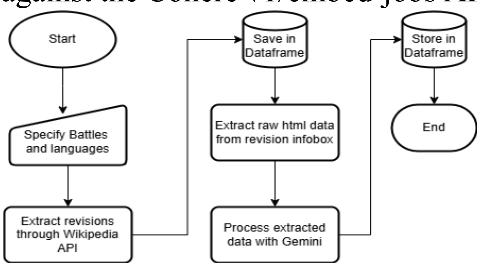


Figure 3: Prototype Version of the Page Content Scraper

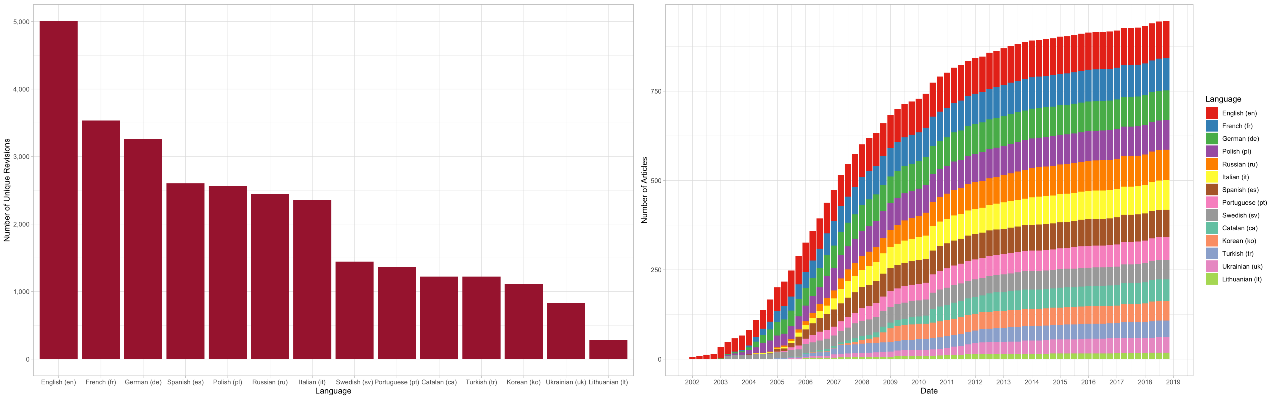
To automate Wikipedia scraping for troop strength and casualty numbers, we use a three-step approach (Figure 3). First, yearly page revisions are collected via the Wikipedia API and saved in a structured data frame. Next, an HTML parser extracts and cleans infobox content for processing. Finally, we deploy the ChatGPT 4o-mini to interpret and translate infoboxes. In a post-processing step we handle inconsistencies across Wikipedia language versions to organize data into structured categories (e.g., troop casualties) with ranges recorded as bounds and thus obtain a harmonize dataset. We deployed the Gemini AI for cross validation. We combine the Gemini AI (flash 1.5 with Google's translation API, and organized data into structured categories (e.g., troop casualties) with ranges recorded as bounds. AI results are validated against manual checks, with discrepancies resolved through Python postprocessing, ensuring accurate, scalable data collection across languages.

The underlying sample includes up to 104 historical battles, out of which 54 (52%) feature full coverage of both at least one winner and one loser. On average, each battle includes 94.4 observations for winners, 80.6 for losers, and approximately 777 uninvolved parties. API information, battle info boxes and cosine similarity measures were combined in the prototype analysis of battle clusters and narrative distances above. Casualty convergence across a second dataset of 46 battles is also available for analysis. Summary statistics can be found in Appendix Table A1.

DATA AND FIRST DESCRIPTIVE ANALYSIS

The left panel in Figure 10 displays the total number of unique Wikipedia article revisions per language edition. This plot highlights variation in editorial activity across language communities, with English, French, and German leading in the number of revisions, suggesting higher attention and editorial engagement in these editions.

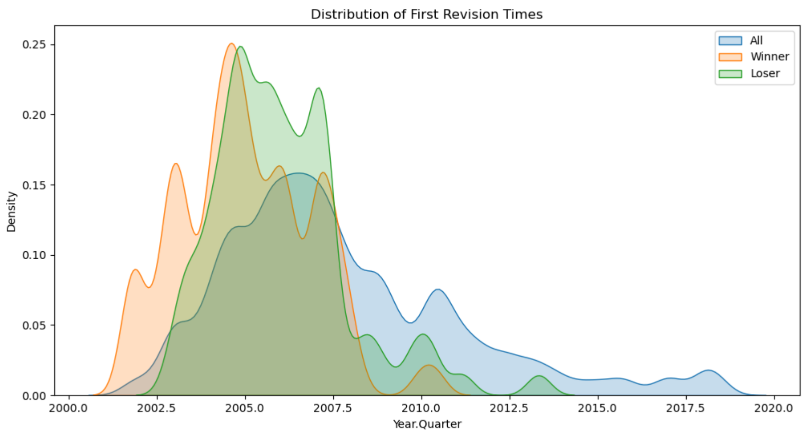
Figure 4: Unique Article Revisions by Languages vs Number of Articles by Language over Time



The right panel in Figure 4 tracks the number of articles per language edition across quarters. This plot captures the expansion and coverage of battle events in Wikipedia over time, as more language editions begin documenting the same conflicts. It also reflects the differential temporal coverage by language, with major editions (English, German, French) generally appearing earlier.

Figure 5 visualizes the distribution of first revision timestamps for Wikipedia articles corresponding to entities involved in historical battles, disaggregated by winner status. The x-axis reflects the year in which an article first appeared on Wikipedia, while the y-axis shows the estimated density based on a kernel density estimator (KDE). We distinguish between three groups: all languages (baseline), languages of losing parties (green, solid line), and languages of winning parties (red, dashed line). The KDEs are computed using only those languages, for which the winner and loser status of the corresponding conflict parties could be clearly identified.

Figure 5: Distribution of First Wikipedia Revisions by Winner Status



Battles associated with winning parties tend to be documented in corresponding languages earlier, while languages associated with losing parties show a rightward-shifted density, suggesting delayed availability of the historical record in its language-specific Wikipedia. This pattern is consistent with the hypothesis that online collaborative platforms may replicate offline visibility asymmetries. Table 1 provides descriptive regressions of how involvement in historical battles influences the timing and magnitude of Wikipedia coverage across three key outcome variables: article creation year, presence in structured battle info boxes, and revision size. Each model controls for three binary indicators: whether the entity was a winner, a loser, or was involved in the battle in any capacity.

Table 1: Impact of Conflict Involvement on Wikipedia Coverage

	(1)	(2)	(3)
	First Revision Year	**First Info box Year**	**Revision Size (bytes)**
Winner	-2.137 ***	-1.030 ***	11108.94 ***
	(0.124)	(0.127)	(340.57)
Loser	-1.127 ***	-0.357 ***	7126.79 ***
	(0.125)	(0.128)	(375.13)
Party involved	-1.089 ***	-0.752 ***	—
	(0.118)	(0.123)	—
Constant	2008.25 ***	2008.31 ***	11920.00 ***
	(0.014)	(0.015)	(155.16)
Observations	64,328	34,774	22,710
Adj. R-squared	0.103	0.065	0.052

Notes: OLS estimates shown with robust standard errors in parentheses. The omitted category consists of uninvolved entities. The dependent variables are: (1) year of first article revision, (2) year of first battle info box appearance, and (3) initial article revision size in bytes. The third model is restricted to articles from battles with complete winner–loser coverage. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Across all three specifications, involvement in a conflict is associated with systematically earlier and more substantive Wikipedia coverage. Winners are, on average, featured over two years earlier than uninvolved actors (Col. 1) and added to battle info boxes roughly one year earlier (Col. 2). Losers are also significantly earlier, though to a lesser extent. In terms of content size, initial revisions for winners are approximately 11,109 bytes longer than for uninvolved entities, while loser-related articles are 7,127 bytes longer (Col. 3). These findings highlight the skewed documentation and visibility of conflict participants—particularly winners—in crowd-sourced historical records.

Together, these results offer preliminary descriptive evidence that Wikipedia's documentation patterns reflect asymmetries in outcome salience, with broader implications for digital memory, historical framing, and knowledge representation in collaborative systems, raising implications for digital historical memory and algorithmic knowledge curation.

QUANTIFYING NARRATIVE DIVERGENCE USING DIFFERENCES IN NUMBERS

We now turn to the quantification of diverging accounts and potential convergence in truths, based on the textual analysis of the articles, and of battle info boxes like the one shown in Figure 1, specifically. For our pilot analysis we use a sample, which includes 99,400 pairwise language observations, drawn from 14 unique languages and spanning up to 104 distinct historical battles (80 after cleaning the data). Figure 5 illustrates the interlanguage availability of structured Battle Box information. Each cell counts how many observations have battle info boxes available for each language pair.

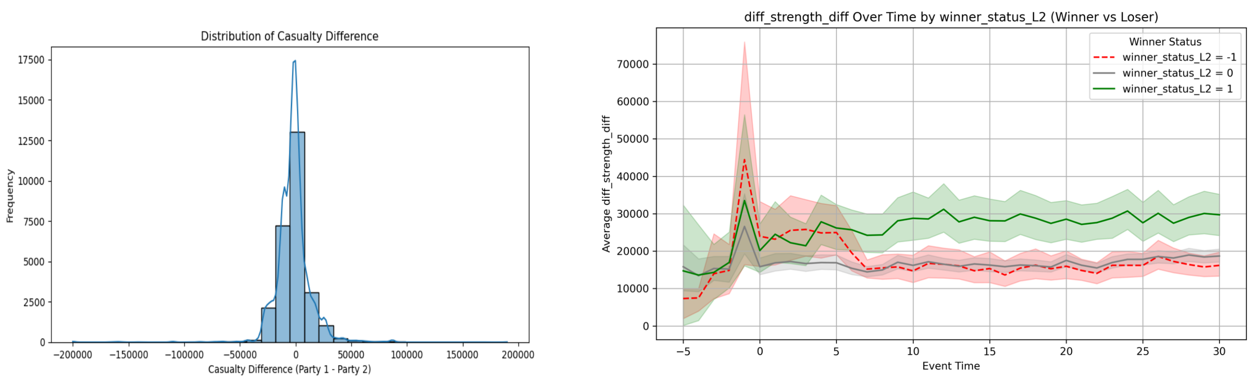
The dependent variables are (1) the absolute difference in reported casualty differentials between the warring parties and (2) the absolute difference in reported troop strength differences (i.e., superiority between two language editions. The first difference derives from comparing casualties of side A to side B in a battle, and our measure then tracks, how these differences differ across languages.

Figure 6: Interlanguage Distribution of Battle Info Box Availability

That the first number – the difference in casualties or troop strength – is naturally non-zero, as it is normally the case that one side fields more troops or loses more troops than the other. However, if both sides agree what happened, then they would both report the same difference in casualties or troop strength, whereas if there is historic (or ongoing) disagreement, then this would be reflected in a non-zero difference in the differences of casualties or troop-strength. For example, if both sides agree that 30,000 soldiers on one side fought 20,000 on the other side, both would report 10,000 as the troop strength difference, such that the difference of these differences is 0 (i.e., no discrepancy). However, if – as in the Battle of Trafalgar – one side reports 30,000 vs. 17,000, while the other side reports 26,000 vs. 18,500, then the difference in these differences would be 5,500, which is the primary numeric measure of narrative disagreement.

Figure 7 shows (in the left panel) the distribution of the difference in reported casualty differences in our sample. The right panel of Figure 7 shows how these differences evolve over time and contrasts it against the placement of a bidirectional language link between the two languages in the observation.

Figure 7: Distribution of Differences in Casualty Differentials and Reported Troop Strength over Time



EVENT STUDY 1: LANGUAGE LINKS and CONFLICTING NUMBERS

The first event study analysis leverages structured info box data which reflects factual claims made within Wikipedia's semi-structured data schema and facilitates visual inference into whether narrative alignment improves after cross-lingual exposure (via hyperlinking). The figure shows a pronounced spike in divergence shortly before language links are introduced, and this spike disappears after the introduction. Furthermore, winner/loser perspectives exhibit different temporal trajectories.

Estimation Strategy and Model Specification: We estimate a high-dimensional fixed effects model to examine how cross-lingual hyperlinking on Wikipedia influences the absolute difference in reported casualty figures between language editions of the same conflict. The dependent variable, *Absolute Casualty Difference*, captures the magnitude of disagreement between two language editions for a given battle and time period, measured in quarters. To identify the temporal dynamics surrounding hyperlink formation, we include a set of event-time indicators ranging from several quarters before to multiple years after the link is established.

The specification includes fixed effects for the battle (i.e., wikidata event ID) and for both source and target languages. These controls account for persistent unobserved heterogeneity across conflicts and language communities. Formally, the regression equation is given by:

$$AbsDiffinCasualtyDiff_{(\mathrm{i}/t)} = \textstyle\sum_k \beta_k \cdot EventTime_{(t=k)} + \alpha_{\mathrm{i}} + \lambda_{l} + \gamma_t + \varepsilon_{(\mathrm{i}/t)}$$

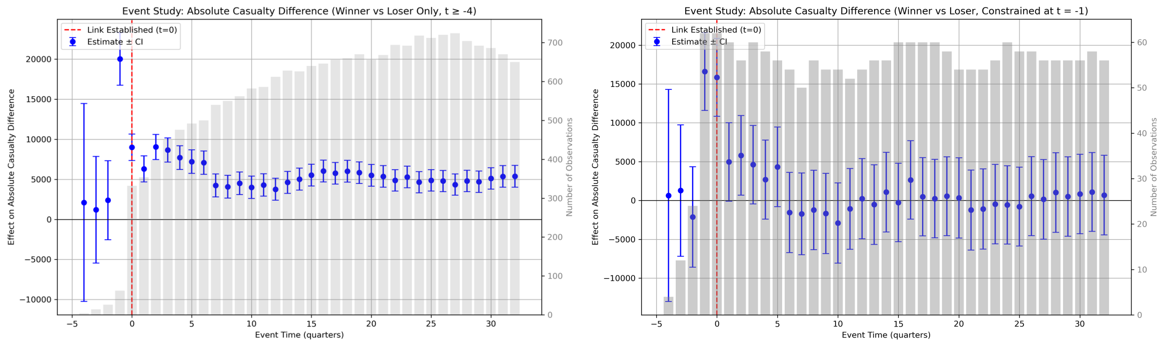
where α_{i}. are battle fixed effects, λ_{l} and γ_t are fixed effects for the time period and the two language editions in the dyad, and ε is an idiosyncratic error term. The model is estimated via OLS with heteroskedasticity-robust standard errors. In an alternative specification we estimate Battle-Language-Language fixed effects.

$$AbsDiffinCasualtyDiff_{(\mathrm{i}/t)} = \textstyle\sum_k \beta_k \cdot EventTime_{(t=k)} + \alpha_{(\mathrm{i}/)} + \gamma_t + \varepsilon_{(\mathrm{i}/t)}$$

Results and Interpretation: The results are presented in Figure 8 below (and in Appendix Table A2). The figure on the left presents an event study of the absolute difference in reported battle casualties across language editions of Wikipedia, from losing and winning parties. Coefficient estimates from a two-way fixed effects regression (battle and language-pair fixed effects) are plotted as blue points with 95% confidence intervals. Estimates for quarters t ∈ [−4, 32] are shown, with the omitted category being t = −1. The secondary axis (right y-axis) displays the number of observations used to estimate each event-time coefficient. The distribution of these counts reveals a sharp growth in coverage around the linking event, with substantially fewer observations available for pre-treatment periods. This imbalance informs the widening of confidence intervals prior to t = 0 and cautions against overinterpreting pre-trend dynamics.

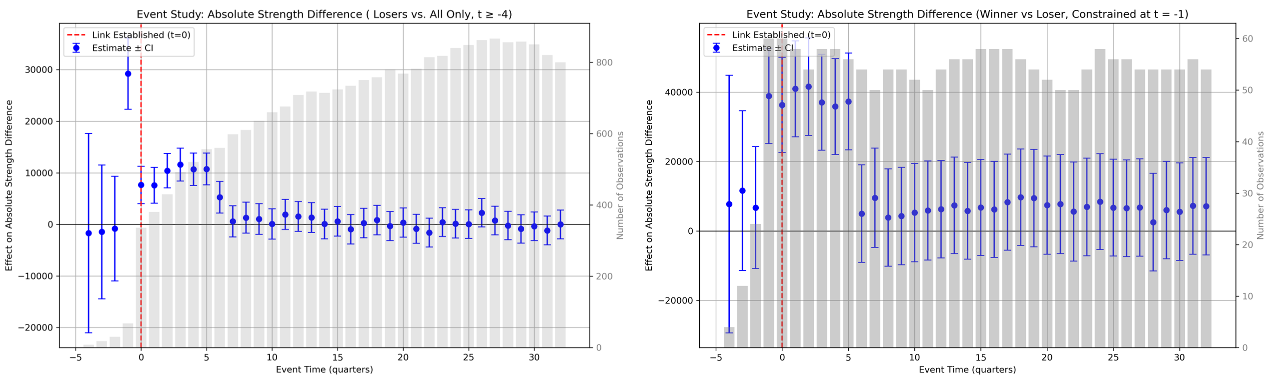
The right panel holds the panel stable by including only observations (Battle-Language-Language) that could be observed in the quarter prior to introducing the language links (i.e. both languages needed to feature an infobox with numbers at least three months prior). The analysis shows an event study design with two way fixed-effects (at the Battle-Language1-Language2 dyad level). The results confirm that the large initial difference in the casualties-differential (of over 15K) quickly decreases to 5000 within 3 months and vanishes after a little over a year.

Figure 8: Differences in Reported Casualty-Differences Across Languages (Losing vs. Other Languages)



Turning from casualty figures to troop strength, we now examine whether Wikipedia's interlingual narratives diverge in their depiction of military capacity. In particular, we analyze the evolution of interlanguage differences in reported troop strength differentials — that is, the extent to which two language editions differ in how asymmetrically they report the number of troops involved in the same battle. As before, we restrict attention to battles where a structural link is established between a "winner" and "loser" page at time t=0 and estimate dynamic treatment effects using the same event-study specification.

Figure 9: Troop Strength Differentials of losing party vs. winners and neutral parties



The results are shown in Figure 9. Interlingual differences in reported troop strength exhibit a sharp and sustained increase following the creation of the interlanguage link. At t=-1, the sample increases and the divergence jumps significantly, implying that these battles are more conflictive. At t=0 the number of battles increases sharply, with patterns, suggesting that the newly observed battles are less conflictive. After that, upon linkage, editors in each language do not appear to revise their pages in a manner that amplifies or reconciles, asymmetries in reported force sizes for over a year.

The effect persists across subsequent quarters and remains well above pre-linkage levels, suggesting that mere exposure to alternative narrative framings may activate editorial responses that reinforce, rather than resolve, contested representations of military strength. However, after 5-6 quarters we see a reduction in the gap, which does suggest convergence.

For the difference in casualty differentials, the pattern of convergence are more pronounced, revealing a sharp and persistent reduction in absolute casualty differences following link formation. There is no evidence of pre-trends: event-time coefficients for the quarters leading up to the link are statistically indistinguishable from zero. Starting at $t = 0$ the establishment of the hyperlink coincides with a large increase of observations, i.e. instances in which both languages under comparison report numbers. The difference is strongly positive and remains statistically significant for over a year but declines in the quarters thereafter. The largest increase is observed one period before the link is introduced suggesting that pre-linking differences were even more substantial. Still, the coefficients for $t = 0$ through $t = 5$ remain in the range of 8,000-10,000 and are significant at the 1% level. This is followed by a decrease to smaller (even though still significant) numbers, consistent with a meaningful decrease in reporting discrepancies. As before we control for sample selection by analyzing the stable sample (in the right panel) and include only observations that are available a quarter prior to the introduction of the language links (t=-1). The convergence pattern is revealed even more clearly in the stable sample.

These findings support the hypothesis that hyperlinking across language editions facilitates informational alignment. By connecting articles on the same conflict, editors may become exposed to alternative narratives or data sources, triggering revisions that reduce interlingual inconsistency. The inclusion of detailed fixed effects ensures that these effects are not confounded by differences in editing intensity or conflict characteristics, strengthening the causal interpretation of the link between hyperlink creation and convergence in reported information.

Outlook and future work: The first descriptive analysis, while promising, highlights limitations of data availability on battle figures. In our main analysis we will deploy our measure of cosine similarity and complement the analysis of open contradictions in numbers with our universally available analysis of narrative convergence or divergence over time. In addition, we will expand the present analysis with more rigorous causal approaches and use pages on celestial bodies as an alternative control group to complement articles about the same topic in uninvolved languages.

QUANTIFYING NARRATIVE DISTANCE USING COSINE SIMILARITY

Our second approach to quantifying narrative distance is based on multilingual embeddings in the Cohere LLM and the computation of cosine distances between two different language versions about the same battle. We compute 1024-dimensional embeddings for each article-language-time observation of the article. To guarantee a stable point of reference, we average across the embeddings of the 2020 version in all languages (over 70), and compute three reference points. The 2020 centroid across all languages, across the 14 core languages in our sample, and across all other languages excluding the 14 focal languages. Then we compute the distance of any given article-language observation at any given point in time to these stable centroids.

Figure 11: Development of Distance Outcomes over Time

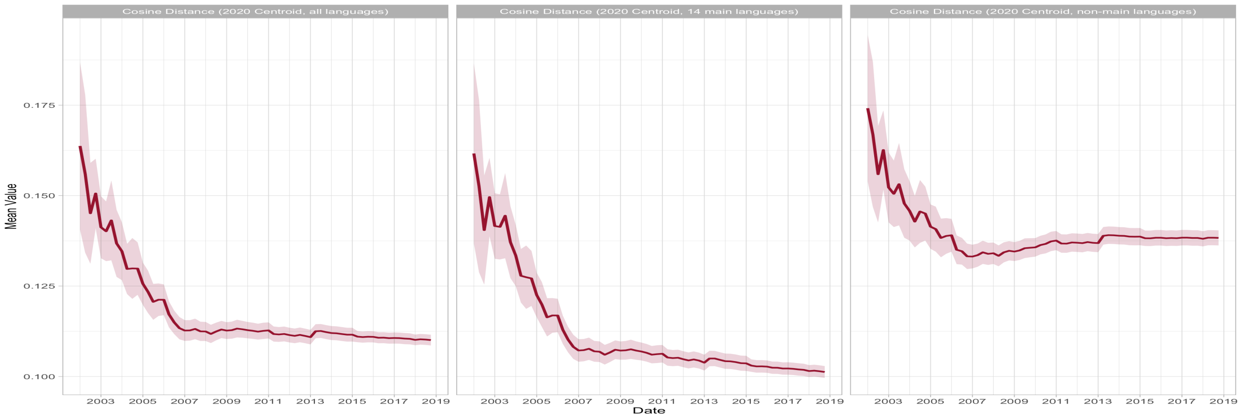


Figure 11 shows the evolution of average cosine distance to different semantic centroids over time: the all-language centroid, the 14-language centroid, and the non-14-language centroid. This plot documents overall convergence of Wikipedia articles across languages, with distances generally decreasing over time, and higher semantic divergence in the non-14 group compared to the 14-language centroid.

Figure 12 tracks the quarterly evolution of three structural article features: number of languages with an article, number of language links, and revision size (in bytes). It shows Wikipedia's increasing content depth and connectivity, with steady growth across all three metrics.

Figure 12: Development of Count Outcomes over Time

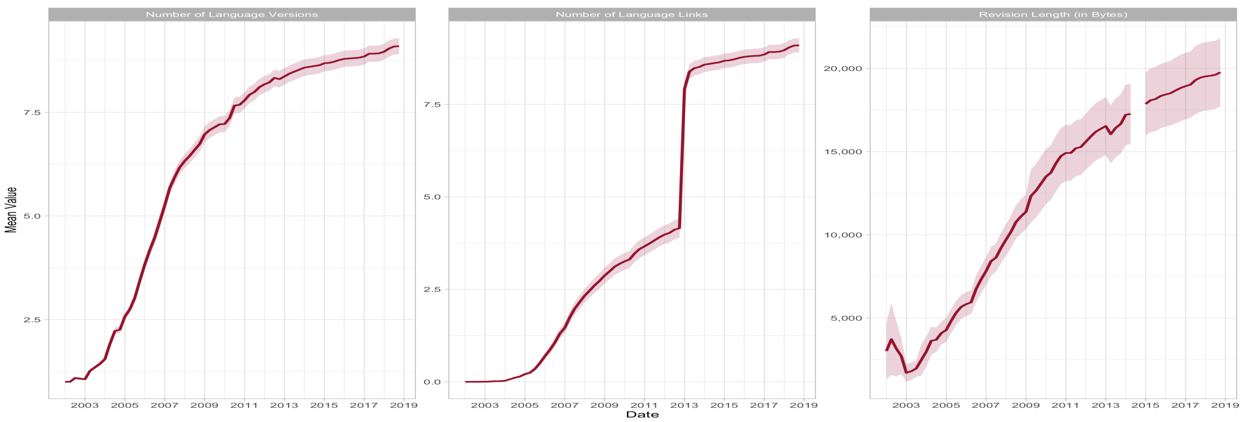


Figure 13 shows the average cosine distance to three semantic centroids (all-language, 14-language, non-14) over time, disaggregated by battle outcome (winner, loser, and other). The plot demonstrates that articles written in the languages of winners tend to be more semantically central (i.e., closer to the centroid), while losers and uninvolved parties begin more distant but converge over time.

Figure 13: Development of Distance Outcomes over Time, by Winner Status

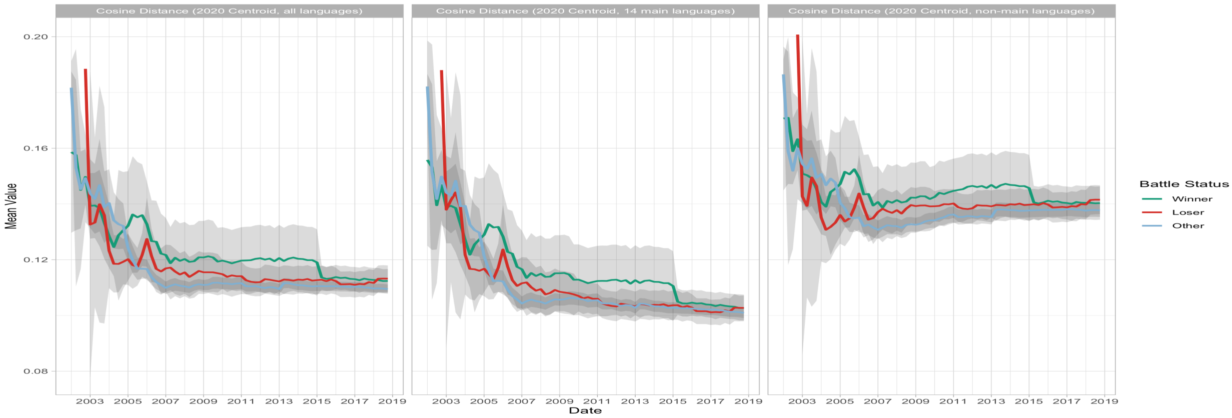
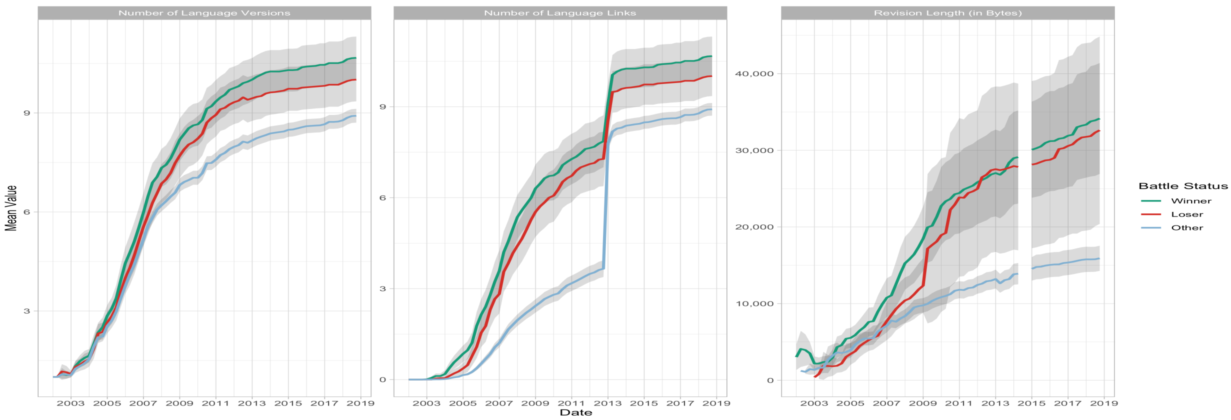


Figure 14 replicates the structure of Figure 12 but disaggregates trends by battle outcome. We can see that winner-language articles are longer, better connected, and more frequently covered across languages—especially in earlier periods—though some convergence is observable over time.

Figure 14: Development of Count Outcomes over Time, by Winner Status



Together, these descriptive statistics document both the continuous growth of Wikipedia over time as new pages are created in new languages and the gradual decline in narrative distances. As this persistent decline coincides with continuous growth in article length and language coverage, we would like to explore whether it is also driven by structural patterns in Wikipedia, and therefore we repeat the event study design using the measures of narrative distance.

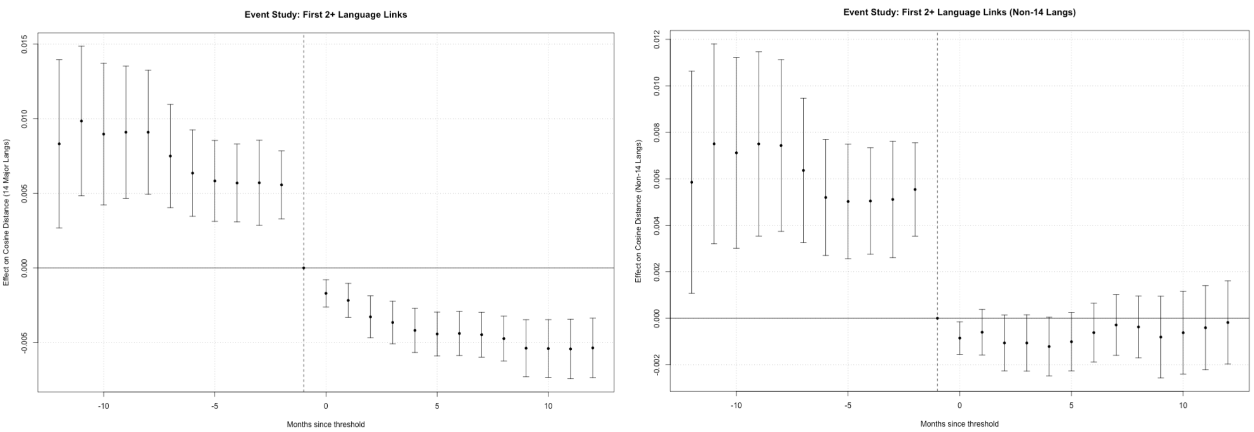
EVENT STUDY 2: LANGUAGE LINKS, INFO BOXES and COSINE DISTANCE

We now proceed to analyzing potential mechanisms that might induce converging narratives. Below we will explore the role of two candidate mechanisms: the placement of interlanguage links and the addition of a standardized formatting scheme to summarize the page: the Battle Infobox.

EVENT 1: Addition of Language Links: In this event study we use the insertion of the second language link as the event. The second language link is the first time the link is likely either to the opponent or to be to a neutral language such as English, or both English and the opposing party. As was hypothesized above that this confrontation with other languages (and alternative versions) could induce a process of reflection and potential consolidation.

As before we present the specification both relative to the 2020 centroids of (i) the 14 focal languages in the study and (ii) of all other languages. The results provide evidence that is in line with the hypothesized pattern and thus underpins the hypothesized convergence in truths as a result of interlinkage and confrontation between the two "opposed" narratives.

Figure 15: Event Study - Addition of the Second Language Link



Notes: Event study estimates of cosine distance to the 14-language (left) and non-14-language (right) centroids, centered on the first occurrence of at least two language links (t=0). Reference period is omitted from estimation. Error bars reflect 95% confidence intervals. Unit fixed effects and clustered standard errors included.

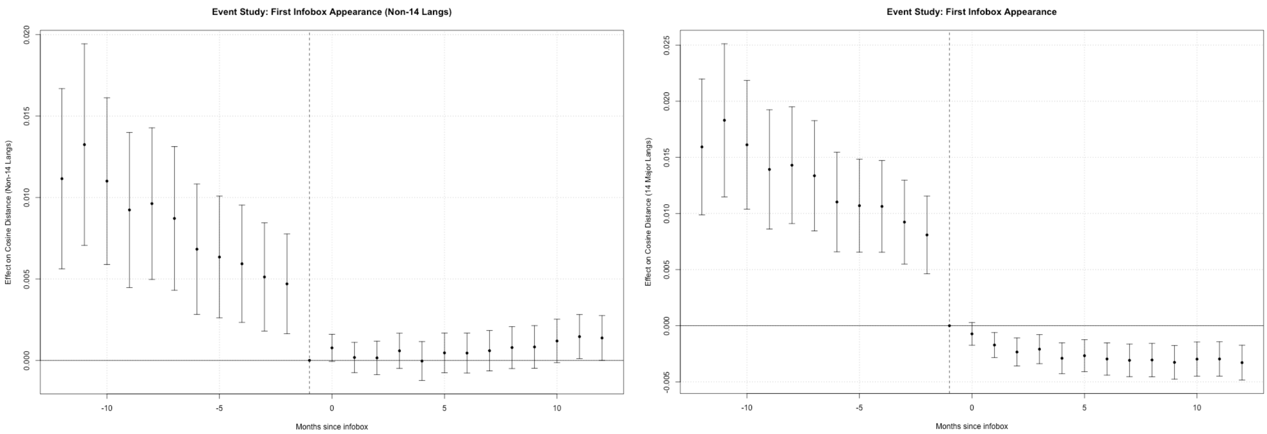
Figure 15 shows event study estimates centered on the first quarter in which an article achieves at least two language links, indicating the beginning of multilingual coverage. The plot structure mirrors that of the infobox event study. The left panel shows cosine distances to the 14-language centroid, while the right panel shows distances to the non-14 centroid. Again, the event time is indexed in quarters, with event_time = 0 denoting the first observed quarter with ≥2 language links and serving as the reference period in the regression. Estimates reflect changes in cosine similarity relative to this threshold-crossing quarter. Confidence intervals at the 95% level are shown, and the dashed vertical line highlights the reference point. Unit fixed effects account for time-invariant characteristics, and standard errors are clustered at the unit level.

These results document a strong decline in cosine distance around the time of introducing the second language link. Regression coefficient estimates are provided in Appendix Table A3.

EVENT 2: Addition of Info boxes: In this event study we use the insertion of the "Infobox" as an event that can reduce narrative divergence. The Infobox enforces a highly prominent and standardized format which was first introduced in English. The standardized format makes it easy for other language readers to read the gist of the article's content, and also makes it easy to compare the article content across different languages. This could induce a confrontation with other languages (and alternative versions) which could induce a process of reflection and potential consolidation.

We present the estimation specification both relative to the 2020 centroid of the 14 focal languages in the study and relative to the 2020 centroid of all other languages (Figure 16). The results provide evidence that is in line with the hypothesized pattern. We find a strong reduction of cosine distance in the article at the time when the infobox is introduced to it. This underpins the hypothesized convergence in truths as a result of interlinkage and confrontation between the two "opposed" narratives.

Figure 16: Event Study: Infobox Appearance



Notes: This figure shows event study estimates of the cosine distance to the 14-language (left) and non-14-language (right) centroids, centered on the first appearance of an infobox (t=0). Coefficients are relative to the reference period and include 95% confidence intervals. Unit fixed effects included; standard errors clustered at the unit level.

DISCUSSION, LIMITATIONS, AND CONCLUSION

This paper investigates whether digital platforms like Wikipedia can facilitate convergence in how different language communities recount historical events. Drawing on a multilingual panel of over 100 battles tracked across more than 50 language editions, we measure both factual and semantic divergence in conflict narratives. Using AI-assisted scraping of structured infobox data and multilingual Large Language Models (LLMs) to embed article texts, we examine two complementary forms of disagreement: reported differences in troop and casualty figures, and narrative distance in article content.

We find that Wikipedia's platform-level features—especially the introduction of interlanguage links and standardized infobox templates—are associated with a measurable reduction in disagreement across language editions. In the case of battle infoboxes, we observe a drop of approximately 4,000–6,000 troops in average cross-lingual differences in reported casualties following the appearance of structured templates. These effects emerge sharply at the point of intervention and persist thereafter. When focusing on the 62 article–language pairs with observable pre-event data, we find that only 43 included casualty information at baseline—but the postintervention convergence is both strong and sustained. While this pattern is consistent with our interpretation, it also reflects an evolving sample and requires further investigation using monthly resolution and continuity-aware sampling.

Cosine distance measures support the same narrative. After embedding articles in a shared cross-lingual semantic space, we find that average semantic distance to a stable 2020 centroid falls by 10–12% following the appearance of language links or infoboxes. Although more abstract than factual differences, this decline indicates growing alignment in tone and content. Importantly, these convergence patterns are not gradual alone—they also exhibit distinctive "spurts" around the moments when visibility and comparability across languages improve.

These findings speak directly to dominant concerns in the literature on online information environments. While most prior work emphasizes how digital platforms create echo chambers and reinforce filter bubbles, our results suggest that under certain conditions, these same platforms can undo them in some contexts. Wikipedia's architecture—designed around transparency, interoperability, and mutual referencing—appears to actively support convergence across national epistemic boundaries, including in domains long shaped by nationalistic historiographies. In this case, the platform does not merely reflect pre-existing filters, but enables their slow and uneven dissolution.

Several limitations remain, and we are addressing them in ongoing work. We are expanding the analysis from quarterly to monthly frequency, improving our ability to detect both gradual and anticipatory dynamics. We are extending the cosine similarity measure to a broader set of languages and battles, with better winner disambiguation to ensure consistent metadata. Most importantly, we are implementing revision-level tracking, which will allow us to directly observe edit patterns, editor spillovers, and content changes in response to cross-lingual visibility. Finally, we are applying this methodology to less contentious domains, such as scientific and sporting events, which will enable difference-in-differences strategies to strengthen causal identification.

In sum, this study contributes new evidence that digital platforms can—under the right institutional design—counteract entrenched narrative fragmentation. In the case of Wikipedia, features that link language editions and highlight inconsistencies appear to catalyze convergence even in deeply rooted historical narratives. As collaborative platforms increasingly shape both public discourse and AI training data, understanding how they can foster epistemic alignment is both timely and consequential.

REFERENCES

Adafre, S. F., & de Rijke, M. (2006). Finding similar sentences across multiple languages in Wikipedia. *Proceedings of the EACL 2006 Workshop on New Text*, 62–69.

Adar, E., Skinner, M., & Weld, D. S. (2009). Information arbitrage across multilingual Wikipedia. *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, 94–103.

Álvarez, G., Oeberst, A., Cress, U., & Ferrari, L. (2020). Linguistic evidence of in-group bias in English and Spanish Wikipedia articles about international conflicts. *Discourse, Context & Media, 35*, 100391. https://doi.org/10.1016/j.dcm.2020.100391

Araujo, A. L. (Ed.). (2013). *Politics of memory: Making slavery visible in the public space.* Routledge.

Assmann, A. (2011). *Cultural memory and Western civilization: Functions, media, archives.* Cambridge University Press.

Assmann, A., & Conrad, S. (Eds.). (2010). *Memory in a global age: Discourses, practices, and trajectories.* Palgrave.

Barberá, P. (2020). Social media, echo chambers, and political polarization. In N. Persily & J. A. Tucker (Eds.), *Social media and democracy: The state of the field, prospects for reform* (p. 34). Cambridge University Press.

Bekerman, Z., & Zembylas, M. (2011). *Teaching contested narratives: Identity and memory in peace education.* Cambridge University Press.

Blanshard, B. (1939). *The Nature of Thought* (Vols. 1–2). George Allen & Unwin.

Bloch, M. (1954). *The historian's craft.* Vintage Books.

Buckley-Zistel, S. (2006). Remembering to forget: Chosen amnesia as a strategy for local coexistence in post-genocide Rwanda. *Africa: Journal of the International African Institute, 76*(2), 131–150.

Carr, E. H. (1961). *What is history?* Macmillan.

Callahan, E. S., & Herring, S. C. (2011). Cultural bias in Wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology, 62*(10), 1899–1915.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230.

Choi, D., Chun, S., Oh, H., Han, J., & Kwon, T. T. (2020). Rumor propagation is amplified by echo chambers in social media. *Scientific Reports, 10*(1), 310.

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences, 118*(9).

Clark, P. (2010). *The Gacaca courts, post-genocide justice and reconciliation in Rwanda: Justice without lawyers.* Cambridge University Press.

Cohere. (n.d.). *Cohere embeddings*. Retrieved 10/12/2024, from https://cohere.com

Cornelißen, C. (2006). Europas Gedächtnislandkarte: Gibt es eine Universalisierung des Erinnerns? In N. Frei (Ed.), *Was heißt und zu welchem Ende studiert man Geschichte des 20. Jahrhunderts?* (pp. 42–49). Wallstein.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.0.

Doxtader, E. (2009). *With faith in the works of words: The beginnings of reconciliation in South Africa.* Michigan State University Press.

Erdmann, M., Nakayama, K., Hara, T., & Nishio, S. (2009). An approach for extracting bilingual terminology from Wikipedia. *Proceedings of the 13th International Conference on Database Systems for Advanced Applications*, 380–392.

Erll, A. (2011). *Memory in culture.* Palgrave Macmillan.

Fallis, D. (2008). Toward an epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology, 59*(10), 1662–1674.

Ford, H., Sen, S., Musicant, D. R., & Miller, N. (2013). Getting to the source: Where does Wikipedia get its information from? *Proceedings of the 9th International Symposium on Open Collaboration*, 9:1–9:10.

Forte, A., & Bruckman, A. (2008). Scaling consensus: Increasing decentralization in Wikipedia governance. *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, 157.

Forte, A., Lampe, C., & Bruckman, A. (2012). Scaling consensus: Collective editing in Wikipedia. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 161–170.

Forte, A., & Lampe, C. (2013). Defining, understanding, and supporting open collaboration: Lessons from the literature. *American Behavioral Scientist, 57*(5), 535–547.

Fu, R., Huang, Y., & Singh, P. V. (2020). Artificial Intelligence and Algorithmic Bias: Source, Detection, Mitigation, and Implications. In *Pushing the Boundaries: Frontiers in Impactful OR/OM Research* (1–0, pp. 39–63). INFORMS. https://doi.org/10.1287/educ.2020.0215

Fu, R., Huang, Y., & Singh, P. V. (2021). Crowds, Lending, Machine, and Bias. *Information Systems Research*, *32*(1), 72–92. https://doi.org/10.1287/isre.2020.0990

Galtung, J. (1996). *Peace by peaceful means: Peace and conflict, development and civilization.* SAGE Publications.

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018). Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference* (pp. 913–922).

Garde-Hansen, J. (2011). *Media and memory.* Edinburgh University Press.

Garton Ash, T. (2011, March 16). Germany can show reborn Arab nations the art of overcoming a difficult past. *The Guardian.*

Greenstein, S., & Zhu, F. (2012). Is Wikipedia biased? *American Economic Review, 102*(3), 343–348. https://doi.org/10.1257/aer.102.3.343

Greenstein, S., & Zhu, F. (2018). Do experts or crowd-based models produce more bias? Evidence from Encyclopædia Britannica and Wikipedia. *MIS Quarterly, 42*(3), 945–959. https://doi.org/10.25300/MISQ/2018/11432

Habermas, J. (1984). *The Theory of Communicative Action: Reason and the Rationalization of Society* (Vol. 1). Beacon Press.

Halatchliyski, I., Kummer, M. E., Slivko, O., Giorgidze, G., & Saam, M. (2016). Centrality and content creation in networks: The case of economic topics on German Wikipedia. *Information Economics and Policy, 36*, 36–52. https://doi.org/10.1016/j.infoecopol.2016.05.001

Halbwachs, M. (1992). *On collective memory.* University of Chicago Press.

Hale, S. A. (2014). Multilinguals and Wikipedia editing. *Proceedings of the 2014 ACM Conference on Web Science*, 99–108.

Hayner, P. B. (2010). *Unspeakable truths: Transitional justice and the challenge of truth commissions.* Routledge.

Hecht, B., & Gergle, D. (2010). The Tower of Babel meets Web 2.0: User-generated content and its applications in a multilingual context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 291–300.

Hinnosaar, M., Hinnosaar, T., Kummer, M., & Slivko, O. (2023). Wikipedia matters. *Journal of Economics & Management Strategy.* https://doi.org/10.1111/jems.12421

Hunt, L. (1989). *The new cultural history.* University of California Press.

James, W. (1907). *Pragmatism: A New Name for Some Old Ways of Thinking.* Longmans, Green, and Co.

Jemielniak, D. (2014). *Common Knowledge? An Ethnography of Wikipedia.* Stanford University Press.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide.* New York University Press.

Kim, S., Park, S., Hale, S. A., Kim, S., Byun, J., & Oh, A. (2016). Understanding editing behaviors in multilingual Wikipedia. *PLOS ONE, 11*(5). https://doi.org/10.1371/journal.pone.0155305

Kirkham, R. L. (2001). *Theories of Truth – A Critical Introduction.* MIT Press.

Knigge, V. (2008). Europäische Erinnerungskultur: Identitätspolitik oder kritisch-kommunikative historische Selbstvergewisserung. In Kulturpolitische Gesellschaft e.V. (Ed.), *kultur.macht.europa – europa.macht.kultur* (pp. 69–82). Klartext.

Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 37–46.

Kummer, M., Slivko, O., & Zhang, X. (2019). Unemployment and digital public goods contribution. *Information Systems Research.* https://doi.org/10.1287/isre.2019.0916

Laitenberger, U., Viete, S., Slivko, O., Kummer, M., Borchert, K., & Hirth, M. (2023). Unemployment and online labor: Evidence from microtasking. *MIS Quarterly.* https://doi.org/10.25300/MISQ/2022/17347

Larson, D. W., & Shevchenko, A. (2010). Status seekers: Chinese and Russian responses to US primacy. *International Security, 34*(4), 63–95.

Lederach, J. P. (2005). *The moral imagination: The art and soul of building peace.* Oxford University Press.

Leggewie, C., & Lang, A. (2011). *Der Kampf um die europäische Erinnerung: Ein Schlachtfeld wird besichtigt.* C.H. Beck.

Lewoniewski, W., Węcel, K., & Abramowicz, W. (2017). Multilingual ranking of Wikipedia articles with quality and popularity assessment in different language editions. *Online Information Review, 41*(4), 478–495.

Lih, A. (2009). *The Wikipedia Revolution: How a Bunch of Nobodies Created the World's Greatest Encyclopedia.* Hyperion.

Liu, S., Yan, X., Li, H., Zhao, T., & Wang, Y. (2018). Cross-lingual entity alignment via attribute-enhanced embedding. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2486–2495.

Luyt, B., & Tan, D. (2010). Improving Wikipedia's accuracy: Is edit age a solution? *Journal of the American Society for Information Science and Technology, 61*(12), 2582–2593.

Maaten, L. van der, & Hinton, G. (2008). Visualizing data using t-SNE. Journal of Machine Learning Research, 9(86), 2579–2605.

Magnus, P. D. (2009). On trusting Wikipedia. *Episteme, 6*(1), 74–90.

Minow, M. (1998). Between vengeance and forgiveness: Facing history after genocide and mass violence. Beacon Press.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2), 203-230.

Miquel-Ribé, M., & Laniado, D. (2018). Wikipedia Cultural Diversity Dataset: A cartography of Wikipedia and Wikimedia projects. *Proceedings of the 14th International Symposium on Open Collaboration*, 1–9.

Montville, J. V. (1990). The healing function in political conflict resolution. Journal of Social Issues, 46(2), 33–53.

Noddings, N., & Brooks, L. (2016). Teaching controversial issues: The case for critical thinking and moral commitment in the social studies. Teachers College Press.

Nora, P. (1996). Realms of memory: Rethinking the French past. Columbia University Press.

Olick, J. K., Vinitzky-Seroussi, V., & Levy, D. (2011). The collective memory reader. Oxford University Press.

Peirce, C. S. (1878). How to make our ideas clear. *Popular Science Monthly, 12*, 286–302.

Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication, 12*(1), 88–113.

Radonic, L. (2010). Krieg um die Erinnerung: Kroatische Vergangenheitspolitik zwischen Revisionismus und europäischen Standards. Campus.

Radonic, L. (2014). Postsozialistische Gedenkmuseen zwischen Opfernarrativen und der 'Europäisierung der Erinnerung.' Jahrbuch für Politik und Geschichte, 5, 85–106.

Rambachan, A., & Roth, J. (2019). An honest approach to parallel trends. *Working paper, Harvard University*.
Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, *90*(5), 2555-2591.
Reagle, J. (2010). *Good Faith Collaboration: The Culture of Wikipedia.* MIT Press.
Regan, P. (2010). *Unsettling the settler within: Indian residential schools, truth telling, and reconciliation in Canada.* UBC Press.
Rescher, N. (1973). *The Coherence Theory of Truth.*
Rescher, N. (1993). *A System of Pragmatic Idealism, Volume II: The Validity of Values.* Princeton University Press.
Ricoeur, P. (2004). *Memory, history, forgetting.* University of Chicago Press.
Rousso, H. (2004). Das Dilemma eines europäischen Gedächtnisses. *Zeithistorische Forschungen, 1*, 363–378.
Rousso, H. (2011). History of memory, policies of the past: What for? In K. H. Jarausch & T. Lindenberger (Eds.), *Conflicted memories: Europeanizing contemporary histories* (pp. 23–38). Berghahn.
Samoilenko, A., Karimi, F., Edler, D., Kunze, T., & Strohmaier, M. (2017). Linguistic neighbourhoods: Explaining cultural borders on Wikipedia through language co-editing activity. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, 217–226.
Simpson, A. (2014). *Mohawk interruptus: Political life across the borders of settler states.* Duke University Press.
Sommer-Sieghart, M. (2006). Historische Ausstellungen als 'Contested Space.' In J. Feichtinger et al. (Eds.), *Schauplatz Kultur – Zentraleuropa: Transdisziplinäre Annäherungen* (pp. 159–166). Studienverlag.
Team, Gemini, Anil, R., Borgeaud, S., Wu, Y., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A. M., Hauth, A., Millican, K., Silver, D., Petrov, S., Johnson, M., Antonoglou, I., Schrittwieser, J., Glaese, A., Chen, J., Pitler, E., ... & Vinyals, O. (2023). Gemini: A family of highly capable multimodal models. arXiv preprint arXiv:2312.11805.
Terren, L. T. L., & Borge-Bravo, R. B. B. R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research, 9*.
Tosh, J. (2015). *The pursuit of history.* Routledge.
Tutu, D. (1999). *No future without forgiveness.* Random House.
Vrandecic, D. (2012). Wikidata: A new platform for collaborative data collection. *Proceedings of the 21st International Conference on World Wide Web*, 1063–1064.
Weizman, E. (Ed.). (2014). *Forensis: The architecture of public truth.* Sternberg Press.
Wilson, R. (2001). *The politics of truth and reconciliation in South Africa: Legitimizing the post-apartheid state.* Cambridge University Press.
Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.
Zehr, H. (2002). *The little book of restorative justice.* Good Books.
Zhang, X. (Michael), & Zhu, F. (2011). Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *American Economic Review, 101*(4), 1601–1615.
Zhu, K., Walker, D., & Muchnik, L. (2020). Content growth and attention contagion in information networks: Addressing information poverty on Wikipedia. *Information Systems Research, 31*(2), 491–509.

APPENDIX

Appendix Table A1: Summary Statistics of Numeric Variables

Variable	min	p10	p25	median	p75	p90	max	mean	sd	n
party_involved	0	0	0	0	0	1	1	0.13	0.34	99009
Is winner	-1	-1	-1	0	1	1	1	0.02	0.97	13261
winner_status	-1	0	0	0	0	0	1	0	0.36	99009
Revision Size (bytes)	21	2342.9	4377	8104.5	15960	33375	391376	14878.25	22263.01	31910
infobox_detected	0	0	1	1	1	1	2	0.83	0.4	42227
has_fully_covered_winner_loser	0	0	0	1	1	1	1	0.52	0.5	99009
battle_fully_covered_for_lang_x	0	0	0	0	1	1	1	0.46	0.5	99009
Strength_Lower_A	0	22	6000	23000	58000	90000	450000	39402.19	47641.36	31786
Strength_Upper_A	0	23	7700	26000	60000	100000	450000	43332.47	52403.89	31774
Total_Lower_A	0	300	880	4300	12466	27000	210000	11008.19	18911.97	28146
Total_Upper_A	0	332	900	4866	14920	30000	220000	12262.22	20669.13	28037
Strength_Lower_B	0	27	7500	27000	60000	103000	460000	45929.82	62285.69	31522
Strength_Upper_B	0	28	9000	30000	66000	120000	700000	51609.31	68927.62	31324
Total_Lower_B	0	627	2808	8000	20000	30000	300000	14206.04	19048.66	29092
Total_Upper_B	0	699	3800	10000	20000	38000	300000	15921.35	20671.75	28958
English (en)	0	0	0	0	1	1	1	0.43	0.49	31910
Spanish (es)	0	0	0	0	1	1	1	0.34	0.47	31910
French (fr)	0	0	0	0	1	1	1	0.41	0.49	31910
German (de)	0	0	0	0	1	1	1	0.39	0.49	31910
Polish (pl)	0	0	0	0	1	1	1	0.38	0.49	31910
Swedish (sv)	0	0	0	0	1	1	1	0.26	0.44	31910
Turkish (tr)	0	0	0	0	0	1	1	0.16	0.37	31910
Russian (ru)	0	0	0	0	1	1	1	0.3	0.46	31910
Italian (it)	0	0	0	0	1	1	1	0.34	0.47	31910
Korean (ko)	0	0	0	0	0	1	1	0.23	0.42	31910
Lithuanian (lt)	0	0	0	0	0	0	1	0.07	0.25	31910
Portuguese (pt)	0	0	0	0	1	1	1	0.26	0.44	31910
Catalan (ca)	0	0	0	0	0	1	1	0.17	0.37	31910
Ukrainian (uk)	0	0	0	0	0	1	1	0.13	0.33	31910
Number of language links	0	0	0	0	0	7	13	1.24	3.12	99009

Variable	Value	n	prop
llm_answer_has_correct_format	NA	64717	0.654
llm_answer_has_correct_format	correct format	34292	0.346

Appendix Table A2: OLS Regression Results of Casualty Differences

Variable	Coef.	Std.Err.	t	P>|t|		
event_time[-4]	2110.79	6299.24	0.335	0.738	-10200.3	14500.3
event_time[-3]	1200.32	3392.83	0.354	0.724	-5449.78	7850.4
event_time[-2]	2399.72	2514.67	0.954	0.340	-2529.15	7328.6
event_time[-1]	20020.00	1672.06	11.971	0.000	16700.0	23300.0
event_time[0]	8995.10	847.19	10.618	0.000	7334.6	10700.0
event_time[1]	6302.98	829.02	7.603	0.000	4678.1	7927.9
event_time[2]	9041.10	798.94	11.316	0.000	7475.1	10600.0
event_time[3]	8659.91	766.90	11.292	0.000	7156.8	10200.0

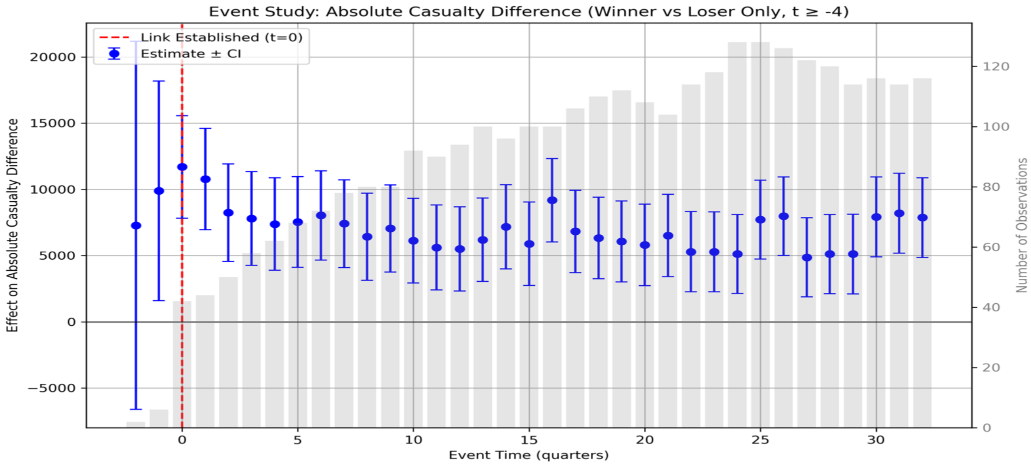
event_time[4]	7713.17	753.82	10.232	0.000	6235.7	9190.7
event_time[5]	7206.33	747.13	9.645	0.000	5741.9	8670.7
event_time[6]	7078.40	743.12	9.525	0.000	5621.9	8534.9
event_time[7]	4231.53	727.37	5.818	0.000	2805.8	5657.2
event_time[8]	4074.23	723.46	5.632	0.000	2656.2	5492.3
event_time[9]	4496.78	719.54	6.250	0.000	3086.5	5907.1
event_time[10]	3997.70	712.53	5.611	0.000	2601.1	5394.3
event_time[11]	4287.16	711.44	6.026	0.000	2892.7	5681.6
event_time[12]	3759.32	703.45	5.344	0.000	2380.5	5138.1

Notes: Standard Errors assume that the covariance matrix of the errors is correctly specified. Number of event-time coefficients before filtering: 66. Number of event-time coefficients in figure: 37.

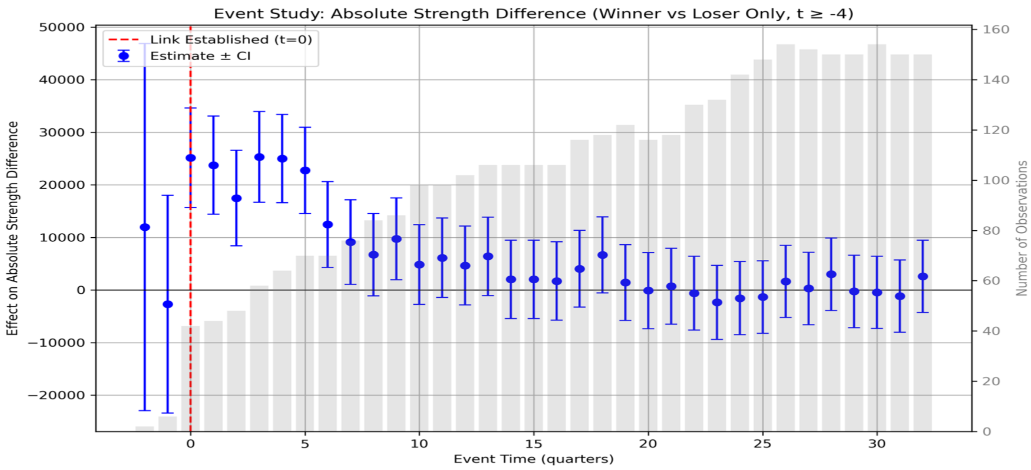
Appendix Table A3: Event Studies: Infobox and Second Language Link

Model	infobox	**(14 languages)**	event	2+ Language Links	**(14 languages)**
Dependent Var.:	cosine_14	cosine_non14	Dependent Var.:	cosine_14	cosine_non14
event_time = -12	0.0159*** (0.0031)	0.0112*** (0.0028)	event_time = -12	0.0083** (0.0029)	0.0059* (0.0024)
event_time = -11	0.0183*** (0.0035)	0.0132*** (0.0032)	event_time = -11	0.0098*** (0.0026)	0.0075*** (0.0022)
event_time = -10	0.0161*** (0.0029)	0.0110*** (0.0026)	event_time = -10	0.0090*** (0.0024)	0.0071*** (0.0021)
event_time = -9	0.0139*** (0.0027)	0.0092*** (0.0024)	event_time = -9	0.0091*** (0.0023)	0.0075*** (0.0020)
event_time = -8	0.0143*** (0.0026)	0.0096*** (0.0024)	event_time = -8	0.0091*** (0.0021)	0.0074*** (0.0019)
event_time = -7	0.0134*** (0.0025)	0.0087*** (0.0022)	event_time = -7	0.0075*** (0.0018)	0.0064*** (0.0016)
event_time = -6	0.0110*** (0.0023)	0.0068*** (0.0020)	event_time = -6	0.0064*** (0.0015)	0.0052*** (0.0013)
event_time = -5	0.0107*** (0.0021)	0.0063*** (0.0019)	event_time = -5	0.0058*** (0.0014)	0.0050*** (0.0013)
event_time = -4	0.0106*** (0.0021)	0.0059** (0.0018)	event_time = -4	0.0057*** (0.0013)	0.0050*** (0.0012)
event_time = -3	0.0092*** (0.0019)	0.0051** (0.0017)	event_time = -3	0.0057*** (0.0015)	0.0051*** (0.0013)
event_time = -2	0.0081*** (0.0018)	0.0047** (0.0016)	event_time = -2	0.0056*** (0.0012)	0.0055*** (0.0010)
event_time = 0	-0.0007 (0.0005)	0.0008. (0.0004)	event_time = 0	-0.0017*** (0.0005)	-0.0009* (0.0004)
event_time = 1	-0.0017** (0.0006)	0.0002 (0.0005)	event_time = 1	-0.0022*** (0.0006)	-0.0006 (0.0005)
event_time = 2	-0.0023*** (0.0006)	0.0002 (0.0005)	event_time = 2	-0.0033*** (0.0007)	-0.0011. (0.0006)
event_time = 3	-0.0021** (0.0007)	0.0006 (0.0006)	event_time = 3	-0.0037*** (0.0007)	-0.0011. (0.0006)
event_time = 4	-0.0029*** (0.0007)	-4.03e-5 (0.0006)	event_time = 4	-0.0042*** (0.0008)	-0.0012. (0.0006)
event_time = 5	-0.0027*** (0.0007)	0.0005 (0.0006)	event_time = 5	-0.0044*** (0.0008)	-0.0010 (0.0006)
event_time = 6	-0.0030*** (0.0007)	0.0004 (0.0006)	event_time = 6	-0.0044*** (0.0008)	-0.0006 (0.0006)
event_time = 7	-0.0031*** (0.0007)	0.0006 (0.0006)	event_time = 7	-0.0045*** (0.0008)	-0.0003 (0.0007)
event_time = 8	-0.0031*** (0.0008)	0.0008 (0.0007)	event_time = 8	-0.0047*** (0.0008)	-0.0004 (0.0007)
event_time = 9	-0.0033*** (0.0008)	0.0008 (0.0007)	event_time = 9	-0.0054*** (0.0010)	-0.0008 (0.0009)
event_time = 10	-0.0030*** (0.0008)	0.0012. (0.0007)	event_time = 10	-0.0054*** (0.0010)	-0.0006 (0.0009)
event_time = 11	-0.0030*** (0.0008)	0.0015* (0.0007)	event_time = 11	-0.0054*** (0.0010)	-0.0004 (0.0009)
event_time = 12	-0.0033*** (0.0008)	0.0014* (0.0007)	event_time = 12	-0.0054*** (0.0010)	-0.0002 (0.0009)
Fixed Effects	Yes	Yes	unit_id	Yes	Yes
S.E.: Clustered	by: unit_id	by: unit_id	S.E.: Clustered	by: unit_id	by: unit_id
Observations	15,988	15,915	Observations	16,119	16,028
R2	0.65877	0.75734	R2	0.78160	0.83350

Appendix Figure A1: Difference in Casualty Differentials between Winners and Losers



Appendix Figure A2: Difference in Troop Strength Differentials between Winners and Losers



Appendix Figure A3: Long Run Effects Analysis