Examining the Relationship Between Core Web Vitals Metrics and Energy Consumption in Mobile Web Apps

Alberto Isotti
2778578
VU Amsterdam
Netherlands
a.isotti@student.vu.nl

ABSTRACT

Context. Core Web Vitals (CWV) is an initiative by Google to provide unified guidance for quality signals that deliver a great user experience on the web. They are a powerful tool for business owners, marketers and web developers to help identify and improve certain aspects of their websites. They offer three stable metrics to measure such quality; Largest Contentful Paint (LCP), First Input Delay (FID) and Cumulative Layout Shift (CLS), and another in pending status Interaction to Next Paint (INP)[1].
Goal. The goal of this project is to investigate whether there is a relationship between these Core Web Vital Metrics and Energy Consumption.
Method. In order to examine the relationship, an experiment is conducted on a Google Pixel 6 and a Pixel 3, where 200 websites are loaded up sequentially, tracking the CWV of each one while monitoring the power consumption of the device.
Results. From the results can be seen a slight correlation between LCP and CLS with energy consumption, with also all the three bandwidth levels playin a role.
Conclusions. This study shows that depending on the characteristics of the web applica- tion, the type of bandwidth and the device, it is possible to correlate some CWV metric values to the energy consumption in mobile devices.

ACM Reference Format:
Alberto Isotti. 2024. Examining the Relationship Between Core Web Vitals Metrics and Energy Consumption in Mobile Web Apps. In *Green Lab 2023/2024 - Vrije Universiteit Amsterdam, September–October, 2023, Amsterdam (The Netherlands).* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Energy efficiency is a key factor when talking about mobile web applications. Devices need good optimization to reduce battery drain. At the moment, it is difficult for a web developer to assess the energy efficiency of the web app during development[2], since measure the energy consumption require specific tools, distinct Frameworks for any type of architecture, and ad-hoc harwdware. Specially with the vast amount of devices present[1].

Websites are everywhere in our daily lives, their complexity has grown exponentially over the years. What was once used to statically host simple text and images has now grown into a dynamic multi-facet system. This huge growth in complexity has not only benefited the end user but also means that loading a web page can involve different processing and fetching steps from multiple servers, since a considerable amount of the content in a web site comes from external sources[3].

Today, the majority of users come from mobile devices which occupy the preponderance of the market [2]. This poses a good reason to analyze their behaviour with websites.

Furthermore, the use of websites on smartphones has also increased due to the growth of **Progressive Web Apps (PWAs)**[3]. PWAs represent another approach to web development in the mobile devices market, since developers can consider not building a native app from scratch, rather combining the functionality of a website with that of a traditional native application. This not only enhances the user experience but also offers the ability to mimic the look of a native app. This means that users could benefit from smoother and faster navigation, while web developers benefit from easy distribution over the Web, without the need to go through mobile app distribution channels, e.g. Google Play, App Store.

Website performance is closely related to energy consumption. When rendering occurs faster, there is less utilization of hardware resources, such as CPU and GPU, which positively impacts overall energy consumption. This is because when fewer components are involved in page rendering and loading operations, it results in a more efficient use of resources, thus reducing overall energy consumption[4].

To help web developers, Google offers **Core Web Vitals (CWV)** [4], a set of metrics used to evaluate website performance.

With this set, web developers can get an idea of what the user experience will be like in order to implement beforehand optimizations to increase the usability of their products.

CWV allows to assess website performance using the metrics that will be explained in Section 3, each with a respective threshold[5]. Developers are particularly interested in those because a good result leads to a better positioning inside the search engine, as Google[6] itself uses them to rank websites.

In order to determine the compliance of a website with CWV, the 75th percentile of page loads needs to meet the targets for each of the four metrics: Largest Contentful Paint (**LCP**), First Input Delay (**FID**), Cumulative Layout Shift (**CLS**). From March 2024, and Interaction to Next Paint (**INP**), a new metric to better assess user

[1]https://gs.statcounter.com/vendor-market-share/mobile

[2]https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet
[3]https://web.dev/progressive-web-apps/
[4]https://w3c.github.io/sustyweb/
[5]https://web.dev/defining-core-web-vitals-thresholds/
[6]https://developers.google.com/search/blog/2020/11/timing-for-page-experience

Green Lab 2023/2024, September–October, 2023, Amsterdam, The Netherlands

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

interaction will replace **FID**, moving to stable status after a period of experimental and pending.

The purpose of this study is to identify the correlation between each of the four CWV metrics and energy consumption in the context of mobile devices. This will help highlight possible adjustments that a web developer can make in advance with respect to the energy consumption and performance, without having to use additional experiments on their web apps.

2 RELATED WORK

Progressive Web Apps. There have been a few studies quantifying performance characteristics of mobile web apps since they have become more pervasive in the last few years. Huber et al. explore this in their study where they measure how much energy Progressive Web Apps (PWAs) consume compared to native mobile applications [5]. They use five different development techniques to implement the same app and compare their respective energy usage in four scenarios against each other as well as native applications. Their aggregated findings show that PWAs consumed at least twice as many joules as native apps in a real-world application test and that this difference is more pronounced on high-end devices than low-end ones. They also observe that the web-browser engine of the PWA had a significant impact, with Google Chrome performing the best in most User Interface (UI) interaction scenarios.

While Huber et al. do test mobile web apps and their impact on the energy footprint of an application, they focus specifically on the energy consumption of UI elements of PWAs and native apps, not whether meeting Google's CWVs correlates with a lower footprint in particular.

Another study regarding PWAs by Malavolta et al.[6] aims to assess the impact that page caching has on the energy footprint as well as quantify its impact on performance in terms of the load time of PWAs. They conclude that while caching has a reasonable impact on page load times, it does not have a serious impact on energy consumption. Therefore Malavolta et al. encourage the use of caching by PWA developers.

This paper explores a similar field since it also studies the energy consumption of mobile web apps. However, Malavolta et al. differ because of their focus on caching and not CWVs.

CWVs. In relation to CWVs, Wehner et al. make use of Google Lighthouse and a subjective user study to find whether CWVs correlate with a more general web quality of experience metric in terms of loading time [7]. They show that the LCP of CWVs does not correlate well with web quality of experience, but page load time and the speed index do correlate positively.

Their study does not consider a possible relation between the energy impact of mobile web apps and focuses on whether CWVs are a useful metric to assess the user experience of a website. While this study does not focus on the energy footprint of (mobile) web apps, their experiment does focus on CWV and questions whether they are an effective metric in the first place.

Energy Efficiency. Chan-Jong-Chu et al. set out to study the correlation between performance scores gathered by Google Lighthouse[7] and the energy consumption of mobile web apps [8]. Lighthouse can provide web performance metrics, some of which are also part of CWVs. These metrics are summarized in a single performance score. Notably, from their results they find that an average to good score significantly correlates with a lower energy footprint compared to a poor performance score in Lighthouse . Their conclusion therefore is to optimise mobile web applications for performance, because this helps with energy consumption as well.

This paper is rather similar in terms of goal and setup, but instead of correlating with CWVs, it investigates a more generic performance score from Google Lighthouse.

Dornauer and Felderer perform a literature study on energy savings and estimate that saving only 5% of energy used could lead to 39 more minutes of battery life [9], emphasizing the importance and usefulness of energy savings with regards to mobile web applications.

Their work provides a comprehensive list of best practices such as which websites to test, sane web browser settings and the number of repetitions used by other papers in the same field as well as good subject selection.

[7]https://developers.google.com/web/tools/lighthouse

3 EXPERIMENT DEFINITION

Core Web Vitals (CWV) provide useful metrics for website developers to evaluate the performance of their products. This permits, in turn, an informed choice of instruments and techniques raising the likelihood of a more fluid and enjoyable user experience. For example performing JavaScript code splitting, loading only the essential script immediately and asynchronously load additional ones, or Minify CSS, JavaScript, and HTML files to remove unnecessary characters and spaces, could lead to some improvement.

There are several stable metrics that are already available, what is missing is information on correlational relationships with energy consumption that would allow at least a connection and partial overview of the phenomena, which aims to encourage developers to take into account energy consumption when they develop software. The purpose of this experiment, thus, is to evaluate the statistical relationship between these core web vitals and the energy consumption of a mobile device when loading mobile web apps.

This requires a sufficient quantity of statistically significant data that will be generated measuring the CWVs of 200 distinct websites using two different Android devices, Google Pixel 6 and Pixel 3, while imposing various bandwidth constraints.

The point of view adopted is that of the developer, and the result of our analysis is aimed at them as their interest in its result is the most prominent among the different possible stakeholders the rest of which will most likely be indirect beneficiaries.

- **RQ1** What is the relationship between Core Web Vitals and Energy Consumption on Android devices?
  - **RQ1.1** What is the correlation of the Largest Contentful Paint metric on Energy Consumption on Android devices?
  - **RQ1.2** What is the correlation of the First Input Delay metric on the Energy Consumption on Android devices?
  - **RQ1.3** What is the correlation of the Interaction to Next Paint metric on the Energy Consumption on Android devices?

- **RQ1.4** What is the correlation of the Cumulative Layout Shift metric on the Energy Consumption on Android devices?
- **RQ2** How does varying available bandwidth influence the interaction between Core Web Vital metrics and energy consumption on Android devices?

To outline a more precise trajectory for the analysis the first question was broken up in four sub-questions that will work as steps to find the answer. This means that the analysis will require the appropriate tools for a multivariate estimation. The second question also necessitates a meticulous investigation since bandwidth may have varying effects, such as moderation or mediation, on the rest of the variables.

4 EXPERIMENT PLANNING

4.1 Subjects Selection

The subject selection starts with the Tranco list [10] of November 2023[8]. The Tranco list combines five providers to generate an averaged list of the 1 million top domains that is more resilient to malicious manipulation. These providers are the Chrome User Experience Report (CrUX), Cloudflare Radar, Farsight, Majestic and Cisco Umbrella. This list used to use Alexa Rank as well, another popular website ranking list, but it was discontinued in 2022 [9].

After obtaining this list, the first 20k web apps have been tested with just one run on the final infrastructure, to cover all levels per each CWV metric (Poor, Need improvement, Goodd), keeping only content rich web applications. From these then have been randomly selected the 200 web apps[9].

4.2 Experimental Variables

The independent variable is the *CWV metric* of a web application. We change the independent variable by testing the 200 different web apps with the three bandwidths. The dependent variable is the *energy consumption* of a web app measured in Joules, during the lifespan of the web page visualization.

4.3 Experimental Hypotheses

In order to answer the two research questions of our investigation, the subsequent hypotheses are developed. To address Research Question 1 (RQ1) and Research Question 2 (RQ2), the researcher formulated the null hypothesis (H0) and the alternative hypothesis (H1).

RQ1.1: correlation of the LCP metric on Energy Consumption:
$H_0 : \beta_{LCP} = 0$
$H_1 : \beta_{LCP} \neq 0$

RQ1.2: correlation of FID metric on Energy Consumption:
$H_0 : \beta_{FID} = 0$
$H_1 : \beta_{FID} \neq 0$

RQ1.3: correlation of INP metric on Energy Consumption:
$H_0 : \beta_{INP} = 0$
$H_1 : \beta_{INP} \neq 0$

RQ1.4: correlation of CLS metric on Energy Consumption:
$H_0 : \beta_{CLS} = 0$
$H_1 : \beta_{CLS} \neq 0$

According to the null hypothesis (H0), the correlation coefficients of the CWV metrics are equal to zero, indicating that CWV has no effect on energy consumption. The alternative hypothesis (H1) suggests that the correlation coefficients of CWV are not equal to zero, which implies that the CWV metrics have significant follow-up on the energy consumption of mobile web applications.

RQ2: bandwidth affects the interaction between CWV metrics and energy consumption in mobile web apps:
$H_0 : \forall i \in \{\text{LCP, FID, INP, CLS}\}, \beta_{\text{bandwidth}\times i} = 0$
$H_1 : \exists i \in \{\text{LCP, FID, INP, CLS}\}, \beta_{\text{bandwidth}\times i} \neq 0$

In this notation, bandwidth: CWV metrics denote the interaction term coefficient between bandwidth and CWV metrics in the Dunn's test after assesd with the Kruskal-Wallis test the existence of odds. According to the null hypothesis (H0), the coefficient of the interaction term is equal to zero, implying that bandwidth has no effect on the relationship between the CWV metrics and the energy consumption. The alternative hypothesis (H1) proposes that the coefficient of the interaction term is not equal to zero, indicating that bandwidth has a significant effect on the relationship between CWV metrics and energy consumption in mobile web applications.

4.4 Experiment Design

The structure of our experiment variables is as follow:

- Core web vitals: It is the main variable in our study, which includes four metrics: LCP, FID, INP, and CLS.
- Websites: are the subjects. Each different website will correspond to a different treatment. Considering that there will be 200 websites tested, this will increase the combinations to test substantially.
- Bandwidth: This will act as a co-factor. Different bandwidths will be tested on all websites, 512 kilobits per second, 2 megabits per second and 50 megabits per second.
- Energy Consumption: the dependent variable.
- Mobile Device: The 200 subjects will be randomly splitted among two devices, Google pixels 6 and Pixel 3, making them a fixed factor for their subject pool.
- Browser: All experiments will be performed on Chrome, making this a fixed factor.

Exploring the interplay of 4 core web vitals metrics, 3 varied bandwidth measurements, and 200 diverse websites leads to a total of 2400 distinct trials. The research aims to conduct correlation analyses, demanding the acquisition of multiple measures for each combination. The strategy involves gathering ten measure for every combination of bandwidth and website, totaling 6000 tests, considering the simultaneous calculation of the four core web vitals during each test session.

[8] https://tranco-list.eu/list/3VP9L/
[9] https://github.com/albertoisotti/Large-Research-Project

4.5 Data Analysis

To obtain an initial comprehension of the nature of the collected data, descriptive statistics are utilized.
In order to conduct the appropriate statistical tests on the data, first step is to determine whether or not they follow a normal distribution. For the purpose of this, qq-plots are employed for the data. Additionally, the Shapiro-Wilk test is utilized to confirm the normality of the data. Some of the strategy tests require normally distributed data for decision making. When assessing hypotheses, the choice is between parametric and non-parametric approaches. Furthermore, it is common practice to calculate informative measures such as the arithmetic mean and standard deviation to offer a comprehensive understanding of the data set.
As a result, the Kruskal-Wallis test is utilized as a non-parametric statistical test for evaluating whether three independent samples originate from populations with identical distributions or for comparing the medians of two groups.
After establishing that significant differences exist between groups with the Kruskal-Wallis test, Dunn's test is used for multiple post-hoc comparisons to determine exactly which pairs of groups differed significantly from each other. Results include z-values for each comparison and associated p-values, as well as p-values adjusted for multiple comparison using Bonferroni adjustment to control the Type I error rate.

5 EXPERIMENT EXECUTION

To measure the energy consumption of CWV, data from the battery of the devices are gathered when loading a web app with the various CWV metrics previously mentioned. This is repeated for 200 web apps, divided on a Google Pixel 6 and a Google Pixel 3.

Both the device are connected to a local WiFi network through a proxy server distinct per device, running on a laptop using Charles proxy [10], which is used to inject some JavaScript code into each web app before it is loaded.

JavaScript injected will load the CWV library[11] with a listener for each metric and an onload event, all of which send a request to Ngrok[12] when triggered. Ngrok offers a tunnel from an HTTPS URL straight to an exposed port on the Laptop, allowing to bypass SSL security checks in the browser. The request is then forwarded to the local NodeJS[13] server which is logging into a CSV file.

Throughout this process, we analyse the energy consumption using Android Runner, running on a Raspberry Pi 3 Model B+[14]. Android Runner is a tool for automatically executing measurement-based experiments on native and web apps running on Android devices.[11], with the BatteryManager plugin, an Android experiment and profiling tool currently maintained by the S2 Group at Vrije Universiteit, Amsterdam. [15]

Experiment execution consists of monitoring each web application using CWV metrics. Before monitoring each web app, Chrome Data is cleared and the script sleeps for 2 minutes to ensure that no previous noise is affecting the output. Then, the monitoring begins by opening the web app in Chrome and after the onload event is triggered the script starts instructing with a tap on the screen, fixed for each website and switching tab, using ADB [16] commands through the Raspberry Pi to perform them. This process is repeated for each web app. During each run the USB charging is disabled.

[10]https://www.charlesproxy.com/
[11]https://github.com/GoogleChrome/web-vitals
[12]https://ngrok.com/
[13]https://nodejs.org/en
[14]https://www.raspberrypi.com/products/raspberry-pi-3-model-b-plus/
[15]https://github.com/S2-group/android-runner

6 RESULT

6.1 Data exploration

When considering a bandwidth of 512kbps, on the Pixel3 in Table 1 tends to show an LCP longer than the Pixel6 in Table 2. This suggests that, under limited network conditions, the Pixel3 may struggle slightly more to make visually significant content available. Instead, looking at the responsiveness of the user interface (assessed through FID and INP), the Pixel6 seems to be less responsive than its contender. Power consumption, a crucial factor for the mobile user experience, shows a slight preference for Pixel6, which tends to be more efficient in this scenario.

Google Pixel 3 Bandwidth of 512 kbps							
Numer of records	Statistic	LCP (ms)	FID (ms)	INP (ms)	CLS	LOAD (ms)	Energy consumption (J)
1000	Mean	2007,38	31,91	98,41	0,059266	4070,73	4070,73
	Median	1543,25	8,50	48,00	0,000000	2128,00	2128,00
	SD	1700,13	78,72	233,19	0,164807	5442,17	5442,17
	Min	237,5	2,9	0	0	2	2
	Max	12892,1	324	2784	1,133691	30000	30000
	CV	0,85	2,47	2,37	2,780820	1,34	1,34

Table 1: Summary statistics for Google Pixel 3 on Bandwidth of 512 kbps

Google Pixel 6 Bandwidth of 512 kbps							
Numer of records	Statistic	LCP (ms)	FID (ms)	INP (ms)	CLS	LOAD (ms)	Energy consumption (J)
1000	Mean	1795,35	65,71	123,54	0,042924	3636,93	3636,93
	Median	1388,70	8,90	64,00	0,000000	1924,00	1924,00
	SD	1606,97	117,19	166,88	0,118102	5131,68	5131,68
	Min	126,8	2,7	0	0	308	308
	Max	15182,4	326,1	2136	1,27902	47394	47394
	CV	0,90	1,78	1,35	2,751449	1,41	1,41

Table 2: Summary statistics for Google Pixel 6 on Bandwidth of 512 kbps

By moving to 2mbps, as shown in Table 3 and 4, we notice that the differences between devices become smaller. LCP times become comparable, suggesting that both devices handle the improved network conditions better. However, responsiveness and power consumption considerations remain consistent with observations made under the 521kbps network. This highlights that while network speed can influence visual loading time, other factors intrinsic to devices play a significant role in their overall responsiveness and power efficiency.

Google Pixel 3 Bandwidth of 2 mbps							
Numer of records	Statistic	LCP (ms)	FID (ms)	INP (ms)	CLS	LOAD (ms)	Energy consumption (J)
1000	Mean	2277,44	35,87	86,62	0,070037	4408,37	4408,37
	Median	1708,40	9,10	56,00	0	2704,00	2704,00
	SD	1861,80	84,34	118,12	0,178115	4105,90	4105,90
	Min	223,00	2,50	0,00	0	1,00	1,00
	Max	18407,70	753,20	1096,00	1,133701	21049,00	21049,00
	CV	0,82	2,35	1,36	2,543151	0,93	0,93

Table 3: Summary statistics for Google Pixel 3 on Bandwidth of 2 mbps

[16]https://developer.android.com/tools/adb

Google Pixel 6 Bandwidth of 2 mbps							
Numer of records	**Statistic**	**LCP (ms)**	**FID (ms)**	**INP (ms)**	**CLS**	**LOAD (ms)**	**Energy consumption (J)**
1000	**Mean**	2213,10	75,22	128,68	0,052462	4470,71	4470,71
	Median	1583,70	9,65	64,00	0,000208	2592,00	2592,00
	SD	2572,89	129,90	165,64	0,132658	4726,66	4726,66
	Min	0,00	2,90	0,00	0	1,00	1,00
	Max	29341,00	1320,20	1560,00	1,030183	31475,00	31475,00
	CV	1,16	1,73	1,29	2,528658	1,06	1,06

Table 4: Summary statistics for Google Pixel 6 on Bandwidth of 2 mbps

With 50mbps, as shown in Table 5 and 6, both devices show notable improvements. Loading times are reduced for both, with the Pixel3 still holding a slight advantage. This improvement reflects the ability of devices to take full advantage of the higher network speeds offered by 50mbps to reduce load times. Responsiveness improves further, and the gap in power consumption between the two devices narrows, indicating that the Pixel6 approaches the Pixel3's performance in terms of efficiency.

Google Pixel 3 Bandwidth of 50 mbps							
Numer of records	**Statistic**	**LCP (ms)**	**FID (ms)**	**INP (ms)**	**CLS**	**LOAD (ms)**	**Energy consumption (J)**
1000	**Mean**	1753,74	36,77	98,62	0,052842	2809,28	2809,28
	Median	1406,30	9,20	56,00	0	2255,00	2255,00
	SD	1198,30	87,14	132,10	0,150228	2244,42	2244,42
	Min	266,00	2,90	0,00	0	3,00	3,00
	Max	7523,50	648,00	1376,00	1,255303	17244,00	17244,00
	CV	0,68	2,37	1,34	2,842974	0,80	0,80

Table 5: Summary statistics for Google Pixel 3 on Bandwidth of 50 mbps

Google Pixel 6 Bandwidth of 50 mbps							
Numer of records	**Statistic**	**LCP (ms)**	**FID (ms)**	**INP (ms)**	**CLS**	**LOAD (ms)**	**Energy consumption (J)**
1000	**Mean**	1569,25	67,73	122,78	0,049772	2524,88	2524,88
	Median	1307,40	9,10	64,00	0	2040,00	2040,00
	SD	1037,93	118,61	160,67	0,128775	1827,93	1827,93
	Min	210,60	2,80	0,00	0	1,00	1,00
	Max	6445,20	325,50	2096,00	1,279020	11557,00	11557,00
	CV	0,66	1,75	1,31	2,587304	0,72	0,72

Table 6: Summary statistics for Google Pixel 6 on Bandwidth of 50 mbps

6.2 Examing for normality

The process of evaluating normality is a common procedure in statistical analysis, employed to determine if a given dataset follows to a normal distribution. There exist multiple methodologies and visual aids that can be employed to assess the normality of a given dataset.

The Shapiro-Wilk test is employed to validate the assumption of normality. As shown in Table 7 for Pixel 3 and Table 8 for the Pixel 6, the four metrics do not follow a normal distribution, as all p-values are lower than 0.05.

	LCP		**FID**		**INP**		**CLS**		**Energy**	
Bandwidth	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**
512 kbps	0.815	**4.56e-32**	0.380	**2.24e-49**	0.615	**4.44e-42**	0.450	**1.77e-47**	0.678	**1.54e-39**
2 mbps	0.887	**4.29e-26**	0.379	**2.21e-49**	0.659	**2.49e-40**	0.399	**6.95e-49**	0.882	**1.22e-26**
50 mbps	0.738	**1.21e-36**	0.343	**2.69e-50**	0.316	**6.01e-51**	0.408	**1.24e-48**	0.724	0.724

Table 7: Shapiro-Wilk test on Pixel 3

	LCP		**FID**		**INP**		**CLS**		**Energy**	
Bandwidth	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**	**W-value**	**P-value**
512 kbps	0.535	**1.33e-45**	0.551	**4.58e-45**	0.659	**5.49e-41**	0.451	**3.53e-48**	0.669	**1.48e-40**
2 mbps	0.880	**2.24e-27**	0.524	**5.99e-46**	0.629	**3.16e-42**	0.446	**2.64e-48**	0.677	**3.16e-40**
50 mbps	0.688	**1.03e-39**	0.516	**3.29e-46**	0.591	**1.13e-43**	0.409	**2.37e-49**	0.833	**2.56e-31**

Table 8: Shapiro-Wilk test on Pixel 6

6.3 Hypothesis Testing

6.3.1 RQ1.1 impact of the LCP on Energy Consumption:

The Tables 9 and 10 illustrates the presence of a statistically significant beta coefficients for LCP in each bandwidths and devices. This finding suggests the existence of a positive relationship between LCP and energy consumption. It can be observed that for each incremental unit of energy consumption, for istance on the Pixel 3 at the bandwidth of 512 kilobits per second, the LCP increase by 68.067.

6.3.2 RQ1.2 impact of the FID on Energy Consumption:

According to the data presented in Table 9 and 10, the analysis reveals a negative beta value for FID in each bandwidths and devices. This statistic indicates a negative correlation between FID and energy consumption. More precisely, the findings indicate that for each incremental unit of energy consumption, there is an anticipated reduction of the FID value, for the Pixel 3 at the bandwidth of 512 kilobits per second, the FID dicreases by 1.2635.

6.3.3 RQ1.3 impact of the INP on Energy Consumption:

According to the findings shown in Table 9 and 10, it can be observed that the impact of INP on energy consumption is only significant at the bandwidth of 50 megabits per second and 512 kilobits per second for both devices. The obtained beta coefficient of 3.4353 for INP on Pixel 3 at bandwidth of 512 kilobits per second indicates a statistically significant and positive association between INP and energy consumption.

6.3.4 RQ1.4 impact of the CLS on Energy Consumption:

According to the findings shown in Table 9 and 10, it can be observed that the impact of CLS on energy consumption is not significant at the bandwidth of 2 megabits per second for Pixel 3 and at the bandwidth of 512 kilobits per second for the Pixel 6. The obtained beta coefficient of 0.0032242 for CLS indicates a statistically significant and positive association with and energy consumption, meaning that for each unit of energy consumption the CLS increse of 0.0032242 in the Pixel 6 at the bandwidth of 2 megabits per second.

CWV Metrics	**Bandwidth 512 kbps**		**Bandwidth 2 mbps**		**Bandwidth 50 mbps**	
	Beta (SE)	**P-value**	**Beta (SE)**	**P-value**	**Beta (SE)**	**P-value**
LCP	68.067 (5.131)	**<2e-16**	89.581 (7.898)	**<2e-16**	15.997 (5.135)	**0.00189**
FID	-1.2635 (0.2549)	**8.4e-07**	-1.7743 (0.3763)	**2.77e-06**	-1.9397 (0.3701)	**1.95e-07**
INP	3.4353 (0.7565)	**6.3e-06**	0.1363 (0.5330)	0.798	2.1128 (0.5648)	**0.000194**
CLS	0.001857 (0.000537)	**0.000568**	0.0014079 (0.0008025)	0.0797	0.0020040 (0.0006437)	**0.0019**

Table 9: Beta coefficients for Pixel 3

CWV Metrics	**Bandwidth 512 kbps**		**Bandwidth 2 mbps**		**Bandwidth 50 mbps**	
	Beta (SE)	**P-value**	**Beta (SE)**	**P-value**	**Beta (SE)**	**P-value**
LCP	65.080 (4.782)	**<2e-16**	134.005 (6.056)	**< 2e-16**	27.589 (4.782)	**1.06e-08**
FID	-2.783 (0.369)	**1.03e-13**	-2.3918 (0.3644)	**8.39e-11**	-5.3871 (0.5291)	<2e-16
INP	-1.1492 (0.5387)	0.0331	-0.3632 (0.4743)	0.444	-1.9371 (0.7498)	**0.00992**
CLS	0.0007451 (0.0003814)	0.0510	0.0032242 (0.0003663)	**< 2e-16**	0.0046473 (0.0005851)	**5.21e-15**

Table 10: Beta coefficients for Pixel 6

6.3.4 RQ2 Bandwidth effect on CWV metrics and energy consumption:

As the assumption that the data comes from a normal distribution was not met, the Kruskal Wallis non-parametric test was performed. The Kruskal Wallis rank based non-parametric test is used to determine if there are statistically significant differences between the bandwidths. For the Pixel 3, the test showed no differences with INP and CLS.

The 11 table offers a detailed overview of how different performance metrics may vary between groups of data representing different bandwidth on the Pixel3. Using a post-hoc analysis using Dunn's test with Bonferroni correction.

For LCP analysis, significant variation in performance between different connection speeds, with all pairs of groups showing statistically significant differences. This suggests that the speed of your internet connection has a marked impact on the loading time of the main page content on the Pixel3.

For FID, differences between groups can still observeed, but with a specific pattern. Performance between 2mbps and 512kbps connections, as well as between 50mbps and 512kbps, differs significantly, indicating that input latency may be affected by connection speed, but not across all speed combinations tested.

With INP, differences between the groups are borderline significant, with the only significant difference emerging between 50mbps and 512kbps. This may suggest that while connection speed affects this metric, the effect is not as pronounced as for LCP or FID.

For CLS, no significant differences emerged between the groups, suggesting that connection speed may not have a direct impact on the visual stability of web pages on the tested device.

Finally, looking at energy consumption, we again see a significant variation related to connection speed, with notable differences between 2 mbps and 512kbps and between 50 mbps and 512kbps. This indicates that connection speed can significantly affect energy efficiency when browsing the web on a Pixel3.

CWV Metrics	2mbps-50mbps			2mbps-512kbps			50mbps-512kbps		
	Z-value	P-value	P-value adjusted (Bonferroni)	Z-value	P-value	P-value adjusted (Bonferroni)	Z-value	P-value	P-value adjusted (Bonferroni)
LCP	5.46	2.357820e-08	7.073461e-08	2.94	1.618535e-03	4.855606e-03	-2.51	5.910684e-03	1.773205e-02
FID	-1.15	0.1250	0.3750	1.99	0.0228	0.0686	3.14	0.0008	0.0024
Energy Consumption	0.97	1.649561e-01	4.948683e-01	4.42	4.921871e-06	1.476561e-05	3.44	2.841734e-04	8.525201e-04

Table 11: Dunn's test results on Pixel 3

The results obtained through Dunn's post-hoc analysis shown in the table 12 provide an examination of the impact of the speed of the Internet connection on the performance and energy consumption of the Pixel6 device. For the Pixel 6, the test showed no differences with INP.

Regarding LCP, the results indicate statistically significant differences between the tested groups. The comparison between 2mbps and 50mbps connections shows the largest difference, with a significant performance improvement in favor of the faster connection. Similarly, the move from 2mbps to 512kbps shows a significant, albeit less dramatic, improvement. The difference between 50mbps and 512kbps, however, is not statistically significant, suggesting that, above a certain threshold, increases in connection speed bring marginal improvements on the LCP metric.

FID is significantly affected by the connection speed when going from 2 mbps to 512 kbps, indicating a reduction in delays in interacting with the web page. Comparisons between 2 mbps and 50 mbps, and between 50 mbps and 512 kbps, show no statistically significant differences, suggesting that for this specific metric, moderate changes in connection speed do not appreciably affect the user experience.

For INP, no significant differences are found between the various connection speed groups. This result implies that the fluidity of interactions after the first page load is not directly influenced by the speed of the Internet connection.

Analyzing CLS, a significant improvement is highlighted going from 2 mbps to 50 mbps, indicating a reduction in unexpected shifts in the page layout. The comparison between 2mbps and 512kbps also shows an improvement, but with a less pronounced difference and only marginally statistically significant. The comparison between 50 mbps and 512kbps shows no significant differences, once again suggesting that speed increases above a certain threshold bring marginal benefits for this metric.

Finally, analysis of energy consumption reveals that connection speed has a significant impact. The jump from 2 mbps to 512 kbps shows the most significant improvement in power efficiency, followed by the comparison between 2 mbps and 50 mbps. The comparison between 50 mbps and 512 kbps, although showing a trend towards improvement, does not reach statistical significance.

CWV Metrics	2mbps-50mbps			2mbps-512kbps			50mbps-512kbps		
	Z-value	P-value	P-value adjusted (Bonferroni)	Z-value	P-value	P-value adjusted (Bonferroni)	Z-value	P-value	P-value adjusted (Bonferroni)
LCP	6.45	5.474483e-11	1.642345e-10	4.35	6.699226e-06	2.009768e-05	-2.09	1.787597e-02	5.362792e-02
FID	1.81	3.446768e-02	0.1034	3.74	9.117099e-05	0.0002	1.92	2.721121e-02	0.0816
CLS	3.65	0.0001	0.0003	2.25	0.0120	0.0362	-1.39	0.0814	0.2444
Energy Consumption	3.26	5.557907e-04	1.667372e-03	5.31	5.461278e-08	1.638383e-07	2.05	2.018133e-02	6.054399e-02

Table 12: Dunn's test results on Pixel 6

Analyzing the interaction between LCP and bandwidth on the amount of energy consumed on the Pixel3 device in Table 13 showed significant results. The effect of LCP on the energy consumed was found to be influenced by the bandwidth used. As bandwidth increases from 512kbps to 2 mbps, and then to 50mbps, the increase in energy associated with an increase in LCP becomes progressively less pronounced. This indicates that for higher bandwidths, the metric has less impact on energy consumption. Although the model suggests that only a fraction of the variance in energy consumed (around 10.47%) can be explained by the variables under consideration, the results remain statistically significant, underlining the importance of considering both web performance metrics and the conditions network in optimizing the energy consumption of devices.

	Estimate	Std. Error	t-value	p-value
LCP:Bandwidth2mbps	-0.0009430	0.0001971	-4.785	1.80e-06
LCP:Bandwidth50mbps	-0.0016264	0.0002526	-6.438	1.41e-10

Table 13: Interaction of LCP, energy consumption among bandwidths on Pixel 3

Analyzing the interaction between LCP and bandwidth on energy consumed, using data from the Pixel6 device in Table 14, provided some interesting insights. While the interaction between LCP and 2mbps bandwidth did not show a statistically significant effect on power consumption, the interaction between LCP and 50mbps bandwidth showed a significant negative impact. This suggests that, contrary to what one might expect, increasing the LCP in high bandwidth conditions (50mbps) tends to reduce energy consumption significantly for the Pixel6 device.

This result may indicate that, for high bandwidths, the Pixel6 device can manage energy resources more efficiently, even with longer loading times. It is important to note that the model explained a larger portion of the variance in energy consumed (approximately 21.86%) than the analysis conducted for the Pixel3 device, suggesting that these factors may have a greater impact on the power consumption of the Pixel6.

	Estimate	Std. Error	t-value	p-value
LCP:Bandwidth2mbps	5.864e-05	1.894e-04	0.310	0.756927
LCP:Bandwidth50mbps	-1.218e-03	2.962e-04	-4.111	4.04e-05

Table 14: Interaction of LCP, energy consumption among bandwidths on Pixel 6

Analysis of the impact of CLS and bandwidth on the energy consumption of Pixel6 devices in Table 15 reveals an interesting pattern. The interaction between CLS and bandwidth shows that as bandwidth increases, the impact of CLS on power consumption varies significantly. For the 2mbps bandwidth, the increase in power consumption associated with an increase in CLS is significantly large, with an average increase of 16.9225 units, indicating that worse layout stability (higher CLS) leads to higher power consumption in this network condition. When the bandwidth is 50mbps, the effect of CLS on energy consumption remains positive but reduces to 7.5444 units, suggesting that with faster connections the impact of low layout stability is less pronounced.

	Estimate	Std. Error	t-value	p-value
CLS:Bandwidth2mbps	16.9225	3.2236	5.249	1.63e-07
CLS:Bandwidth50mbps	7.5444	3.2670	2.309	0.0210

Table 15: Interaction of CLS, energy consumption among bandwidths on Pixel 6

7 DISCUSSION

The investigation explored the correlation between Core Web Vitals metrics and energy consumption on Android devices, focusing specifically on the Pixel 3 and Pixel 6 models. Through beta coefficient analysis and the Kruskal-Wallis test, accompanied by post-hoc analysis, it is possible to see how various web performance metrics influence energy consumption in relation to internet connection speed.

LCP demonstrated a positive correlation with energy consumption, highlighting how longer loading times are associated with greater energy expenditure. This is consistent with the hypothesis that more extensive loading processes require greater computational resources, reflected in an increase in energy consumption. The significant variation in LCP in relation to the bandwidth further confirms the importance of the network in the overall performance and energy consumption of devices.

In contrast, FID shows a negative correlation with energy consumption. This suggests that improvements in page responsiveness can help reduce energy consumption, possibly through reduced processing requirements due to more efficient handling of user input.

INP showed a positive correlation with energy consumption only under specific bandwidth conditions, suggesting that the impact of this metric on energy consumption can vary significantly depending on connection quality. This indicates that optimizing user interactions may be particularly relevant in contexts of limited connectivity.

CLS showed a positive correlation with energy consumption in less extensive circumstances than other metrics, implying that visual stability may influence energy consumption but to a lesser extent than factors such as content loading time.

Significant variations in energy consumption related to connection speed highlight how efficient connectivity management can represent a crucial aspect in optimizing energy use during web browsing. This is especially true for transitions from low to medium connection speeds, where the most significant improvement in energy efficiency is seen.

The analysis of individual web sites highlights how the increase in bandwidth can have diversified effects on the performance of web pages and their energy consumption. While some sites show improvements in performance metrics resulting in an increase in energy consumption, others manage to maintain or improve performance without a significant impact on energy consumption, highlighting the importance of site-specific optimizations, this is something that web developers can take a look to find the right compromise between performance and energy consumption. LCP turned out to be the most important metric to care about for this purpose.

8 THREATS TO VALIDITY

8.1 Internal Validity

One possible internal threat may come from running the experiment in an inconsistent environment. Therefore, when running the experiment, the Android device as well as the Raspberry Pi and the Laptop running the proxy have as few background processes running as possible. The browser's data are cleared before each new measure is taken to mitigate against caching. The proxy helps to minimize network interference.

Furthermore, web apps are in continuous development. Their performance in terms of CWV can improve or deteriorate over time. For this experiment, all measurements were made in January and February 2024.

Lastly, the load of the web server the web app is running on might also affect the resulting CWVs. To ensure reliability, each web application is tested 10 times.

8.2 External Validity

Firstly, a threat might arise from how the use of a proxy differs compared to the real world. There will be significantly more latency if the test is run using real-world servers compared to the proxy.

This is not an issue for this experiment since the proxy is used consistently across the whole experiment. Therefore, the difference in latency and therefore in CWV metrics will also be consistent. A correlation between CWV metrics and energy consumption can still be found, regardless of ordinary web performance.

Secondly, the web browser used to run the experiment might not be representative of the target population since the experiment is run on a single web browser, Google Chrome. Different browsers, such as Mozilla Firefox, use a different rendering engine and could therefore use different amounts of power. Google Chrome browser has a market share of 64% at the time of the experiment[17] and is therefore representative for the majority of mobile web users.

Also, the way of how the interaction is simulated, in a real scenario just one tap on the screen is not representative of the typical human way of browsing mobile applications.

8.3 Construct Validity

A possible threat in terms of construct validity comes from not properly defining what energy consumption is and how it is measured. To mitigate against this, GQM was used to help design the experiment. Throughout the experiment, important factors such as the device model, software versions and network conditions are either noted down or dealt with.

8.4 Conclusion Validity

Firstly, low statistical power is a possible threat. To combat this the experiment uses 200 different websites and runs them 10 times each in order to have a sufficient number of data points. If the results are not statistically significant, the conclusion must reflect this.

Secondly, the experiment does not assume a normal distribution. Section 6.2 checks whether the samples are normally distributed.

Lastly, the non-parametric Kruskal-Wallis test combined with the Dunn's test are used if the results are not normally distributed to be able to draw conclusions from the data.

9 CONCLUSIONS

This research explored the correlation between Core Web Vitals and energy consumption in the context of mobile applications, highlighting how Largest Contentful Paint (LCP) and Cumulative Layout Shift (CLS) factors show a positive correlation with energy consumption, regardless by bandwidth. In contrast, First Input Delay (FID) appears to correlate negatively with energy consumption. Significantly, the results confirm that bandwidth plays a determining role in energy consumption. While Interaction to Next Paint (INP) showed no correlation with energy consumption.

These findings offer mobile web application developers valuable insights into selecting tools geared toward sustainable digital practices, potential cost savings from reducing energy consumption, and extending device battery life. However, it is critical to recognize the limitations of this study, including limited resources and time that affected data collection and randomization of variables, limiting the applicability of the findings.

To advance our understanding of this topic, it is suggested to extend the research to a wider range of devices and collect a greater volume of data. This would not only refine estimates of the relationship between Core Web Vitals and energy consumption but also expand the applicability of the results. Further investigations could also explore the impact of other variables, such as the use of different network technologies or varying app workloads, to offer a more holistic understanding of how development practices influence the energy sustainability of mobile applications .

[17] https://gs.statcounter.com/browser-market-share/mobile/worldwide

REFERENCES

[1] P. Walton, "Web vitals | articles," May 2023. [Online]. Available: https://web.dev/articles/vitals

[2] K. Chan-Jong-Chu, T. Islam, M. M. Exposito, S. Sheombar, C. Valladares, O. Philippot, E. M. Grua, and I. Malavolta, "Investigating the correlation between performance scores and energy consumption of mobile web apps," in *Proceedings of the 24th International Conference on Evaluation and Assessment in Software Engineering*, ser. EASE '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 190–199. [Online]. Available: https://doi.org/10.1145/3383219.3383239

[3] M. Butkiewicz, H. V. Madhyastha, and V. Sekar, "Understanding website complexity: measurements, metrics, and implications." in *Internet Measurement Conference*, P. Thiran and W. Willinger, Eds. ACM, 2011, pp. 313–328. [Online]. Available: http://dblp.uni-trier.de/db/conf/imc/imc2011.html#ButkiewiczMS11

[4] G. LLC, "Core web vitals," 2023. [Online]. Available: https://web.dev/vitals/

[5] S. Huber, L. Demetz, and M. Felderer, "A comparative study on the energy consumption of progressive web apps," *Information Systems*, vol. 108, p. 102017, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306437922000230

[6] I. Malavolta, K. Chinnappan, L. Jasmontas, S. Gupta, and K. A. K. Soltany, "Evaluating the impact of caching on the energy consumption and performance of progressive web apps." in *MOBILESoft@ICSE*, D. Lo, L. Mariani, and A. Mesbah, Eds. ACM, 2020, pp. 109–119. [Online]. Available: http://dblp.uni-trier.de/db/conf/icse/mobilesoft2020.html#MalavoltaCJGS20

[7] N. Wehner, M. Amir, M. Seufert, R. Schatz, and T. Hoßfeld, "A vital improvement? relating google's core web vitals to actual web qoe," in *14th International Conference on Quality of Multimedia Experience (QoMEX), Best Paper Award*, Lippstadt, Germany, 9 2022.

[8] K. Chan-Jong-Chu, T. Islam, M. M. Exposito, S. Sheombar, C. Valladares, O. Philippot, E. M. Grua, and I. Malavolta, "Investigating the correlation between performance scores and energy consumption of mobile web apps." in *EASE*, J. Li, L. Jaccheri, T. Dingsøyr, and R. Chitchyan, Eds. ACM, 2020, pp. 190–199. [Online]. Available: http://dblp.uni-trier.de/db/conf/ease/ease2020.html#Chan-Jong-ChuIE20

[9] B. Dornauer and M. Felderer, "Energy-saving strategies for mobile web apps and their measurement: Results from a decade of research," *2023 IEEE/ACM 10th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*, pp. 75–86, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:257921560

[10] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, "Tranco: A research-oriented top sites ranking hardened against manipulation," in *Proceedings of the 26th Annual Network and Distributed System Security Symposium*, ser. NDSS 2019, Feb. 2019.

[11] I. Malavolta, E. M. Grua, C.-Y. Lam, R. de Vries, F. Tan, E. Zielinski, M. Peters, and L. Kaandorp, "A framework for the automatic execution of measurement-based experiments on android devices," in *35th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW '20)*. ACM, 2020, pp. 61–66. [Online]. Available: http://www.ivanomalavolta.com/files/papers/A_Mobile_2020.pdf