Persuading users into verifying online fake news

University of Oulu
Faculty of Information
Processing Science
Master's Thesis
Md Hasan Shahriar
Date 17.06.2019

Abstract

Checking authenticity of fake news before sharing online can reduce spread of misinformation. But fact-checking requires cognitive and psychological effort, which people are often not willing to give. Some fact-checking methods might even be counterproductive, entrenching people into their deeply held beliefs. Numerous online fact-checking services have emerged recently which verify false claims to address the issue. While these services are quite efficient technologically, they seriously overlook human behavioral factors associated with fake news.

Persuasive systems have been proven successful in attitudinal and behavioral changes, which could be applied here as behavioral interventions for fact-checking. A review of current fact-checking services showed that they significantly lack persuasive features, resulting in a passive and linear user experience. Findings from cognitive science and persuasion literature paved way for development of a fact-checking mobile application that would encourage users into regular fact-checking.

Qualitative and quantitative evaluation of the artifact showed promise of persuasion in combating fake news. Social support persuasive features were found most effective, followed by tunnelling and self-monitoring. Implications of these findings and future research directions are discussed.

Keywords
Fake news, fact-checking, persuasion, PSD model, behavioral intervention

Supervisor
Professor and Dean Harri Oinas-Kukkonen

Foreword

My heartiest thanks and gratitude to Professor Harri Oinas-Kukkonen for guiding me in my first academic research endeavour. Although small at scale, I always wanted to work on solving an important problem of the world through scientific research, which led me to choose fake news as a topic. This thesis may not have been the solution to the complex problem, but it surely sheds some new lights on the topic.

I also want to thank my parents and family for their endless support and love, throughout my studies and career. I am grateful for the opportunities you blessed me with. Thanks to my friends and peers in Oulu for keeping me motivated even in the winter.

Hasan Shahriar

Oulu, Finland

17.06.2018

Contents

1. Introduction

Fake news has been as an important research agenda for human-computer interaction discipline for some time (Hecht et al. 2017). In the backdrop of 2016 U.S. presidential election, it was clear that researchers have a greater role to play to tackle issues that emerge and scale unprecedentedly by leveraging the deep and complex relationship between human and computers (Hecht et al. 2017; Brown et al. 2017). A concrete and action-based research agenda on misinformation has been called by many researchers, who emphasized to focus on tangible outcomes rather than purely descriptive frameworks. One intervention recommended by many experts was to build fact-checking tools that encourage users to evaluate authenticity of news before sharing (Hecht et al. 2017; Lazer et al. 2018).

Indeed, numerous news verification services have been developed recently to address the issue in academia and industry. However, efficacy of these services has been questioned by Fernandez et al. (2018), Brandtzaeg et al. (2017), Brandtzaeg et al. (2018) and Karduni (2019). For example, Facebook's much-hyped 'disputed flags' effort to identify false news was shut down when study found visual flags might actually have a backfire effect (Lyons, 2017). Several fact-checking tools developed in academia were technologically marvel, but were unknown to even journalists who need such tools on daily basis (Brandtzaeg et al., 2018). Fernandez and Alani (2018) found that independent fact-checkers such as Snopes or PolitiFact lack usability and treat users as 'passive consumers' without actively involving them.

Based on recent literature, it can be argued that inefficiency of most fact-checking tools is caused by overlooking social and cognitive factors associated with fake news. The psychology of fake news stem from our collective cognition to stick with pre-existing beliefs (Lazer et al., 2018) and tendency to seek biased information (Lazer et al., 2017). Thus, pure technological applications without social and cognitive interventions are not sufficient in empowering and educating individuals about fake news. It's up to the users at the end of the day to critically evaluate information before sharing (Torres, Gerhart and Negahban, 2018).

Persuasive systems could be used here to encourage individuals into fact-checking. Systems built on persuasive strategies have been widely applied in healthcare to change people's behavior or attitude. Such systems also helped users to achieve certain goals and encouraged them into physical activities (Orji and Moffat, 2018; Forget et al., 2008; Wunsch et al. 2015). These systems create an interactive experience for users by applying psychological and behavioral cues (Fog, 1998; Sundar et al., 2010). With advancement of Web 2.0, most domains have seen adaptation of interactive systems that deeply engage users. The domain of fact-checking, in comparison, remained quite mechanical unfortunately. Services in this domain significantly lack the persuasive features that make modern applications so interactive (Table 3). Many fact-checking services still operate as bare-bones websites, whereas fake news managed to take advantage of social web to become an epidemic.

The aim of this thesis was to design a fact-checking behavioral intervention that changes users' behavior towards fake news. The intervention was expected to encourage users into regular fact-checking, and form the habit of authenticity verification before sharing news. The research was guided by the following research question:

RQ: How can we design a persuasive mobile application that encourages users to check authenticity of online fake news before sharing?

The main research question was supported by an additional research question: *Which cognitive and behavioral factors influence people into fact-checking?*

To address the research question, design science research approach as proposed by Hevner et al. (2004) was followed to build and evaluate a software artifact. Following Peffers et al.'s (2007) suggestions, approach of the research was problem-centric. Problem domain and state of the art were extensively studied to build the theoretical grounds, which was then implemented in an artifact using software engineering method.

Requirements of the final application were identified from prior scholarly work on cognitive science. Several theories from cognition literature were found strongly relevant in fact-checking context, which could potentially help individuals to overcome the psychological biases of fake news. Oh and Sundar's (2015) work on interactivity was found relevant in terms of increasing cognitive effort, while Lewandowsky et al. (2012) proposed behavioral interventions based on cognitive cues. Khan and Idris (2019) analysed fact-checking behavior with Theory of Reasoned Action and found that motivating people to verify news before sharing is a major way to reduce misinformation spread. One of the most comprehensive mitigation strategies was proposed by Endsley (2018) based on her Situational Awareness (SA) theory, which emphasized heavily on cognitive engineering methods. She proposed factors such information presentation, confidence, integration and mental models to develop effective fact-checking solutions which could override our cognitive vulnerabilities.

To turn these theoretical findings into a persuasive artifact, Persuasive System Design (PSD) model was used (Oinas-Kukkonen and Harjumaa, 2009). The model provides a set of design principles to build and evaluate persuasive systems. Four categories of persuasion are described in the model, out of which most relevant features were selected for the artifact based on the thesis's scope. A high-fidelity iOS application consisting of functional and mockup features was built based on the final design principles. After development, it was evaluated to measure usability, desirability and persuasiveness. Results from the tests confirmed several findings from cognitive literature, along with promise of persuasion as a potential strategy to build fact-checking systems.

The research is novel in a way that it tries to connect several threads of prior findings to develop a human-centered solution to the problem of fake news. The focus was on overall usability and persuasiveness of the application, as opposed to the most fact-checking tools which depends on authoritative methods to deal with the complex socio-technical problem. Such nuanced approach has been an important agenda in fake news literature, and this is the first study to connect persuasion with fact-checking as per author's knowledge. Findings from this study could have implications for similar future research.

The thesis is structured into seven chapters. The first chapter provided an overview of the research and explained motivation behind the topic. Second chapter is background and relevant work, where theoretical background and findings from literature are discussed. Four major threads of research are explained in detail here, with an analysis of current fact-checking services. The third chapter explained the design science research method adopted for this study. The fourth chapter illustrated the design and development of the application. Chapter five follows up with the evaluation process of the artifact, with detail about three results and their analysis. The sixth chapter discussed overall findings of the study. The seventh chapter provided concluding remarks with guidance for future studies.

2. Background and Relevant Work

This section discusses background of the research and relevant literature. Related works in the area has been divided into five sections: fake news, fact-checking services, cognitive elements of fact-checking, persuasive technologies and paradoxes.

2.1 Fake news

Amid all the recent media attention about fake news, it seems like we have forgotten one simple fact: fake news is not a new phenomenon. Its history goes as back as the invention of printing press in 1439 (Soll, 2018). At the very first opportunity to share and distribute human knowledge in a collective form, some people misused the technology to spread deceptive news and misinformation.

Researchers, especially from communication and social science areas, emphasized on this historical notion to conceptualize fake news. Tandoc, Lim and Ling (2018) gave a classic example of widespread misinformation from 1938, when a drama by H.G. Wells was broadcasted on radio just for entertainment, but caused huge panic as people interpreted it as factual news. Allcott and Gentzkow (2017) gave another historical example from 1835 when an American newspaper published a series of articles claiming life was discovered on moon. They also provided a list of hoaxes that has been believed by significant amount of American population over the past half-century.

The definition and characteristics of fake news have evolved over time. Today, fake news is usually associated with misinformation in internet and social media (Allcott and Gentzkow, 2017; Tandoc et al., 2018). While there is no agreed-upon definition of the term ‘fake news’ (Shu et al., 2017; Pierri and Ceri, 2019), a general and widely acceptable definition has been given by Allcott and Gentzkow (2017), where authors described fake news as news articles that are “intentionally and verifiably false, and could mislead readers”. The term has been used by researchers from different disciplines to portray the problem of online misinformation. Tandoc et al. (2018) reviewed 34 articles on the topic and found that the term is used to define a broad of range of concepts such as news satire, news fabrication, news parody, photo manipulation and propaganda. Rubin et al. (2015) treated any type of deceptive news as fake news i.e. as hoaxes, fabrications and satires. Aligning with Rubin et al. (2015), Allcott and Gentzkow (2017), Tandoc et al. (2018), Karduni (2019) and Sharma et al. (2019), I adopt the broad definition of fake news for this research. In this article, the term fake news will be used to define any false information online regardless of the intention behind it.

If fake news has been around for a long-time, what causes its recent notoriety? Experts almost unanimously agree that the reason is 2016 U.S. presidential election. They alleged that influence of fake news might have been pivotal in election of Donald Trump as U.S. president (Allcott and Gentzknow, 2017; Lazer et al. 2018). By leveraging the news consumption behavior of internet and social media, fake news has reached a new magnitude with this election, manipulating election results with fake news stories circulated through websites, Facebook posts and tweets. Similarly, fake news on social media impacted the 2016 Brexit referendum and 2017 French national election (Pierri and Ceri, 2019). It caused public chaos during natural disasters in Japan and U.S., affected stock market price and instigated public violence like the infamous ‘Pizzagate’ shootout (Sharma et al., 2019). A survey by BuzzFeed found that 75% of Americans were fooled by fake news headlines (Tandoc, Zim and Lim, 2018). Looking at its growing magnitude and diversity, it’s clear that digital media has turned fake news into one of the major

global risks of our time (The Global Risks Report, 2018). The journalistic niche that started from Gutenberg's wooden press in the middle-age found the perfect channel of social media to exploit human vulnerabilities, at scale. It's almost the 'plague' of digital age (Shandwick and Tate, 2016).

The unique characteristics of modern internet make it a breeding ground for fake news (Zhang and Ghorbani, 2019). The reason it managed to reach such magnitude was because social platforms like Facebook dramatically lowered the access barriers to information, creating a new power structure where any individual can distribute large volume of information without facing gatekeeping (Lazer et al., 2017; Allcott and Gentzknow, 2017). This has enhanced the spread of low quality news that defers traditional journalistic norms and third-party fact-checking (Pierri and Ceri, 2019). Consuming news from social media is also faster and cheaper than traditional news media, which caused it to replace television as the major source of news (Shu et al., 2017). At the same time, general trust towards traditional mass media like TV and newspapers declined due to many socio-political reasons (Lazer et al., 2018). Furthermore, presence of numerous malicious agents such as automated or semi-automated bots on social platforms make them vulnerable to rapid dissemination of misinformation (Shao et al., 2017). Many people take advantage of these factors due to a number of economic and ideological motives (Allcott and Gentzknow, 2017), creating large volume of misinformation and disseminating them on social platforms.

The meteoric rise of fake news, however, goes beyond technical aspects of social media. Fake news exploits the human vulnerabilities to maximize its impact and make it indistinguishable from real news (Shu et al., 2017; Sharma et al., 2019; Lazer et al. 2017). First of all, humans do not have any natural ability to discern real information from false ones (Shu et al., 2017). This makes us prone to naïve realism and confirmation bias. The former makes people believe that only their perceptions of reality are true, while the ones who disagree are biased or irrational. Confirmation bias, or selective exposure means we are only seeking information that confirms our existing beliefs. Another relevant psychological theory is normative social influence, which says that individuals prefer 'socially safe' options to be consistent with the norms of their community, which is important for personal identity and self-esteem (Shu et al., 2017; Sharma et al., 2019). In modern news consumption systems, these theories over-ride our rational decision-making mechanism and incline us towards immediate social rewards, even if that involves sharing fake news.

These psychological theories take a collective shape in social media and online platforms, creating an effect called echo chamber (Flaxman et al., 2016; Shu et al. 2017). Recent studies show that these echo chambers are one of the primary drivers of fake news diffusion (Shu et al. 2017). In these chambers, individuals only exchange information that are aligned with their existing ideologies. This results in social homophily, where people only make connections based on similar ideologies. On the other hand, social platforms use personalized algorithms to promote contents that align with users' preferences (Shu et al. 2017; Karduni, 2019; Sharma et al., 2019). These factors combined create echo chambers, also known as filter bubbles, where users do not get exposed to opposing viewpoints. Lazer et al. (2018) described it as a situation where attitudinal polarization is amplified and tolerance towards opposition is minimized. The reasons fake news gets high traction within these polarized communities are two psychological phenomena called social credibility and frequency heuristic (Shu et al., 2017). Social credibility means people will believe an information if others from their community do so, and frequency heuristic suggests that people can believe even a false information if they are repeatedly exposed to it. According to Sundar (2008), bandwagon heuristics become

more prevalent in these social groups, and people continue to consume and share same information over and over again. The result is isolated information bubbles with little information about outside world, which is a ripe environment for birth and dissemination of fake news (Shu et al. 2017; Karduni, 2019; Sharma et al., 2019).

Structure or elements of fake news has been studied by Karduni (2019), Zhang and Ghorbani (2019), Shu et al. (2017), Sharma et al. (2019). Karduni (2019) and Sharma et al. (2019) broke down fake news into three major components: source, content and consumer. Source is the outlet through which fake news is consumed i.e. fake news websites, social bots, individuals who share/redistribute false information. Content broadly includes the physical content of fake news such as textual and visual components, manipulative information etc. The content aspect has been further broken down by Shu et al. (2017), who proposed that fake news content usually contain publisher name, headline, body text and photo/video. These features together try to manipulate consumers with linguistic and visual cues. Consumers are individuals who are exposed to fake news and takes several decisions based on the news such as trusting the news and sharing it with peers (Shu et al., 2017). Using these elements as framework. Karduni (2019) further studied their inter-relationships in different dimensions. He argued that the complexity of consumer-content relationship is the main leeway through which fake news manage to spread, as this relationship is affected by multiple psychological, cognitive and social processes. Apart from Karduni's (2019) three elements frame, Shu et al. (2017) and Zhang et al. (2019) also proposed a fourth element of fake news, which is the social context. Social context refers to the dissemination process of fake news, and how it drives social engagement. This context provides useful auxiliary information to understand impact of fake news (Shu et al., 2017).

Detection and mitigation techniques of fake news is another thread that gained popularity among scholars in last two years. Fernandez and Alani (2018), Mosinzova et al. (2019), Pierri and Ceri (2019), Zhang and Ghorbani (2019), Karduni (2019), Sharma et al. (2019) reviewed mitigation literature and characterized them based on different criteria. Pierri and Ceri (2019) and Mosinzova et al. (2019) categorized them under three major categories: content-based, context-based and content-context (hybrid) based. Content-based methods depend on meta-information contained in the body of fake news. Different types of lexical, semantic and computational analysis has been applied to detect veracity of a news content. Context-based methods rely on user behavior and their interaction with fake news articles to prevent dissemination. There has been different approaches in this method such as identifying fake news based on Facebook likes, retweets, crowdsourcing, web traffic analysis. Hybrids methods combine elements of content and context-based methods to provide better accuracy (Mosinzova et al. 2019).

Shu et al. (2017) presented several open issues for future direction of fake news research, which were followed as a guideline by later scholars. They proposed four approaches to fight fake news: data-oriented, feature-oriented, model-oriented and application-oriented. Among them, application-oriented research was proposed as a method that hinders diffusion of fake news through proactive or reactive social intervention. Sharma et al. (2019) called for new type of intervention strategies that take educational or gamified approach. Zhang and Ghorbani (2019) proposed several directions to build more efficient detection systems through unsupervised learning models, real-time classification, early prediction. Mosinzova et al. (2019) called for a language-independent approach so that non-English speaking nations can also use the system.

Many scholars also pointed out the lack of interdisciplinary approach in current literature. Karduni (2019) noted that existing mitigation research do not incorporate findings from

other disciplines which are pivotal in combating fake news. Notions like cognitive bias, social exposure and normative influence are overlooked in existing computational tools, as well as the empirical finding that direct fact-checking is counterproductive in many cases where it contradicts deeply hold beliefs (Lyons, 2017; Lazer et al., 2018). Over-reliance on technological aspects made current prevention methods limited and unscalable (Lazer et al., 2017; Lazer et al., 2018). Lazer et al. (2018) also proposed to develop interventions that can empower and motivate individuals to check authenticity of fake news before sharing. Actively engaging users in fake news identification process was the most important direction for future research according to Fernandez and Alani (2018). Overall, a number of very recent papers urged to take an interdisciplinary approach to combat the problem of misinformation, and emphasized that social and cognitive aspects are just as pivotal as technical aspect.

2.2 Fact-checking services

Fact-checking services or news verification services analyse and determine accuracy of online claims and provide guidance for users to verify those claims (Brandtzaeg et al., 2018). It's one of the major mitigation techniques to combat fake news (Fernandez and Alani, 2018). Duke Reporters Lab conducted a survey in 2017 and identified 114 independent fact-checking services (Stenchel, 2017). Individuals or organizations operate these services and usually publish their opinion through websites. Apart from these services, several fact-checking tools have been developed in recent years by academia and industry. These tools usually support the process of news verification through technological solutions such algorithms, search engines and web plugins. While it's difficult to distinguish these groups, one major difference between them is that services usually verify facts through journalistic approach, and tools take a more technological approach. Individuals behind services can use these tools to fasten and improve their verification process. In this thesis, the term 'fact-checking services' will mean all types of services, tools and websites associated with news verification.

Websites like Snopes, PolitiFact and FactCheck.org are early examples of fact-checking services that rely on manual detection by professional organizations or journalists (Brandtzaeg et al. 2018). They usually take a journalistic approach towards news verification, and a set of editors determine truthfulness of claims by manually checking facts, keywords and context. Reports are published on their websites as news pieces, often with ratings depending on authenticity of a certain claim. Some services like PolitiFact provides APIs for third-party applications to access their fact-checking database (Politifact). Snopes has emerged as a high-quality service and was profiled by major global publishers such as CNN, New York Times, Fortune (Zhang and Ghorbani, 2019). Many countries also have localized fact-checking services run by non-partisan groups or organizations. While these services received accolades from different bodies for their contribution to fake news detection, their impact in reducing spread of misinformation has been limited. This is due to the manual and time-consuming nature of journalistic fact-checking, which often cannot match the meteoric speed of news sharing (Hassan et al., 2017).

Automated fact-checking, on the other hand, provides more comprehensive methods for fake news verification and correction. Most automated tools try to predict the chance of a particular news piece being purposefully deceptive (Karduni, 2019). ClaimBuster is a recent example of such system that leverages machine learning and natural language processing to predict if sentences in an article are 'check-worthy' (Hassan et al., 2017). It's an end-to-end system consisting of several components like Claim Monitor, Claim Matcher, Claim Spotter. Similar attempts have been made to fact-check through

knowledge graphs, statistical scoring and source cross-matching (Karduni, 2019). Another group of researchers analysed lexical and semantic features of news content to distinguish false information. Bourgonje et al. (2017) used logistic regression classifier to compare consistency of news headlines with body content. Rubin et al. (2016) detected satirical cues in news content with high accuracy by using an SVM algorithm. Potthast et al. (2017) found that extremist and hyper-partisan world views in information can be distinguished from factual news just based on writing style. Few researchers also tried to identify misinformation based on visual information. Jin et al. (2017) extracted several features of microblog images such as clarity, coherence and similarity, and applied classification models to determine their credibility. Gupta et al. (2013) studied fake photos distributed during a hurricane and achieved high accuracy in distinguishing real images from fake ones. Diffusion pattern of fake news has been studied by some scholars who conducted temporal analysis of news propagation. Wu and Liu (2018) used such method to characterize fake news based on social network structures. Another wave of research tried to identify different sources of misinformation and wanted to minimize their reach. Social bots, a common source of fake news, got significant attention of scholars here, which was countered with different graph-based, feature-based and crowd-based bot detection methods (Karduni, 2019). In an early study, Diakopoulus et al. (2012) went beyond social bots and developed two classifiers to help journalists verify sources of fake news.

While most computational and manual verification services lacked an engaging user experience, some researchers tried to develop interactive and user-friendly fact-checking systems. TweetCred was developed by Gupta et al. (2014) to verify content of Twitter in real-time. It provides a visual rating system based on machine learning to determine credibility of each tweet. It can be used as a web plugin and users will see ratings in Twitter web page real-time. Twittertrails is another service built on Twitter that provides more interactive web-based fact verification (Finn et al. 2014). Users can explore the level of dissemination and skepticism for a particular claim, and see their temporal nature and visibility level. through interactive visualizations. RumorLens was one of the first tools that combined computation with interactivity to visualize the complex process of news propagation (Resnick et al., 2014). A similar tool is RumorFlow which provides visual analytics of fake news and shows visualization of rumor topic, word cloud, relationship between topic and rumor (Dang et al., 2016). Hoaxy, an acclaimed tool developed by Shao and his colleagues (2016), automatically combines information stream from different social media accounts and show how misinformation is spread in their networks. The tool also provides a comprehensive search engine to search ongoing hoaxes and claims. Its dashboard shows timeline of fake news spread and visualization of accounts from where the news is spread. Some studies went beyond interactive tools and proposed new interaction techniques for news verification. Pourghomi et al. (2017) took into account some psychological aspects of fake news and proposed a ‘right-click authentication’ approach in contrast to direct fact-checking. Nguyen et al. (2018) designed a human-AI interface where users are actively involved in news verification process.

Social media and large internet companies also developed their own fact-checking services to address the problem of misinformation. Facebook took one of the early initiatives in December 2016 and started adding ‘Disputed’ tags to the posts which were found fabricated by third-party fact-checkers (Clayton et al., 2019). This approach, however, was turned down just after one year as research found that authoritative fact-checking has backfire effect and strong confrontation only reinforces people’s existing beliefs (Lyons, 2017; Clayton et al., 2019). Facebook replaced fact-checking badges with ‘Related Articles’ feature that simply tries to give more context of news, which was more

effective than previous approach. Google published several tools recently under 'Google News Initiative' to fight misinformation and promote quality journalism (Funke, 2018). WhatsApp launched a fact-checking service in India ahead of national election where users can assess credibility of claims through text messages (Lomas, 2018).

An overview of different types of fact-checking services is provided below. These services were developed by different bodies and take distinguished approaches towards news verification. Thus, analysing them provides an illustrative picture of fact-checking approach from that domain. For example, services developed by academic researchers focused more on technical aspects like data visualization while independent fact-checkers focused on journalistic aspects like exhaustive reports.

Table 1: A general overview of different types of fact-checking services

Service/ tool	**Approach**	**Medium**	**Key features**
Snopes	Journalistic	News website	Extensive reports, credibility rating
PolitiFact	Journalistic	News website, mobile app	Extensive reports, virtual 'truth-meters', gamified app, API
ClaimBuster	Automated/ end-to-end	Web tool	Real-time claim monitoring, sentence analysis, fact-check repository, knowledge base, API, Slackbot
Hoaxy	Automated/ visualization	Web tool	Search engine, interactive dashboard, social media accounts visualization
TwitterTrails	Automated/ visualization	Web tool	Real-time tweet monitoring, timeline and spread visualization, related images, interactive dashboard
Facebook	Mixed	Social media, mobile apps	'Disputed' flags (discontinued), 'Related Articles' to provide background information of news

While many of these fact-checking tools and services received acclaim for accuracy and credibility, their usefulness and trustworthiness to mass users has been questioned (Lazer et al., 2018; Brandtzaeg et al., 2018). Empirical evidences on their efficacy are still very limited (Karduni, 2019). Many experts raised questions like who would authenticate the fact-checkers, and how they would distinguish between opinions and facts. More relevant to the topic of this thesis is the criticism of their usefulness. Online fact-checking services are information systems, and their success depends on their perceived usefulness (Brandtzaeg et al., 2018). This usefulness depends on whether users perceive them as necessary and eligible to discern false information from factual news (Davis, 1989). Current body of research suggests that usefulness of existing fact-checking services are, at best, mixed (Lazer et al., 2018). The most common barrier to their usefulness, as found by Nguyen et al. (2018), Fernandez and Alani (2018) and Lazer et al. (2017) and Karduni

(2019) is lack of a nuanced approach that considers behavioral and cognitive aspects of fact-checking.

2.3 Cognitive elements of fact-checking

Fact verification by individuals is a complex cognitive process. There is a consensus among psychologists that simply presenting correct information to people is unlikely to change their existing belief or opinion (Fernandez and Alani, 2018). The backfire effect of such confrontation has been confirmed by many studies (Nyhan and Reifler, 2010; Lazer et al., 2018). Unfortunately, current fact-checking services and tools overlook this cognitive complexity and just strive to achieve technical accuracy (Nguyen et al., 2018; Karduni, 2019). This absence of human factors in news verification tools made most of them somewhat inefficient. Most of these tools contradict with the ultimate goal of fact-checking, which is to reduce the effect of misinformation on users (Karduni, 2019).

Fact-checking tools communicate through interactive media like websites and applications. Interactive media try to influence users through different forms of content (Jensen, 1997; Xu and Sundar, 2014). One form of interactivity is medium-based interactivity or modality interactivity. Modality interactivity means the availability of different types of tools and cues on interfaces that maximize accessibility. Interfaces with high number of modalities can influence users' attitude and behavior towards a content (Xu and Sundar, 2014). For example, a website with substantial clicking and scrolling features would be perceived as more interactive than websites that merely shows static images. These high modality interfaces can achieve greater cognitive absorption, which means users will be paying more attention to the content. Cognitive absorption, or user engagement in simple words, has been conceptualized as a mental state where users are emotionally and cognitively involved with a particular task (Xu and Sundar, 2014). While engaged in a task, users fully invest cognitive resources to process incoming information and make judgement based on that. Cognitive absorption and message elaboration are two aspects that can maximize user involvement with an interface, leading to persuasion (Xu and Sundar, 2014). In other words, user engagement is the key to achieve such level of persuasion that could lead to behavioral change (Oh and Sundar, 2015).

Interactive media from most domains significantly evolved towards persuasion by applying user engagement techniques in last couple decades (Sundar et al., 2010). Health applications evolved from informative websites to interactive mobiles apps that are highly personalized and customizable. Online shopping experience has been enhanced by visual and psychological cues such as 3D view, rating system. Tourism websites heavily use persuasive features to retain users in their sites (Ibrahim et al., 2013). Social networking websites like Facebook harnessed persuasive strategies to keep billions of users engaged in their system (Fogg et al., 2008).

Comparing with these domains, fact-checking media remained mostly primitive. Current fact-checking services not only lack interactive features, they also significantly lack basic level of user involvement. As Fernandez and Alani (2018) noted, current fact-checking tools disengage users by treating them as passive consumers rather than co-creators and detectors of fake news. On the other hand, fake news manages to engage readers by leveraging from low cognitive effort (Osatuyi et al., 2018) or 'mental shortcuts' (Sundar et al., 2007). To battle this problem, cognitive and behavioral interventions are required that would take a nuanced and interactive approach towards news verification (Lewandowsky et al., 2012). Sundar et al. (2007) also found that perceived credibility of online news is not just restricted to the news content. News interfaces that use psychological cues are likely to receive higher credibility and user engagement, even in

situations where the task needs extra cognitive effort (Sundar et al., 2007). These findings have implications for fact-checking behavior. It means if users could be deeply engaged in fact-checking services with interactive features, they might be willing to spend cognitive energy to evaluate misinformation.

Another interesting observation from Table 1 is that most of the independent fact-checkers don't have smartphone applications, which means smartphone users only reach them through mobile browsers. This significantly restricts the scope of user engagement. Imagine health applications where you could only read about healthy habits but could not track your physical activities.

Several attempts have been made to connect psychology theories with fact-checking behavior. In an early study, Lewandowsky et al. (2012) provided a comprehensive framework grounded in cognitive psychology to design effective interventions. The authors emphasized on role of individuals in reducing misinformation, and proposed that skeptic attitude and prior warnings are keys to prevent its cognitive influence. Based on Theory of Reasoned Action, Khan and Idris (2019) concluded that individuals should be at the centre of the efforts while dealing with spread of misinformation, and should verify news by themselves before sharing. Kumar and Geethakumari (2014) developed a generic framework based on cognitive decision-making process and proved that having users to make conscious decisions is a way to fight misinformation. Based on theoretical research on epistemology of testimony and trust, Torres et al. (2018) found that perception of news author, news sharers' network and intention to share news influences news verification behaviors. Fernandez and Alani (2018) suggested to closely engage users in the misinformation detection and verification process as a fundamental step of fact-checking. Understanding users' motivation and personality traits are also key factors (Chen and Sin, 2013).

Table 2: Constructs relevant to fact-checking, extracted from cognitive literature

No.	Construct	Relevant theory	Author
C1	High-modality interface	Interactivity Cognitive absorption	Oh and Sundar (2015)
C2	Mechanism to cognitively engage users		
C3	Inclusion/responsibility of users	Theory of Reasoned Action	Khan and Idris (2019)
C4	Skepticism about news source	Worldview Skepticism	Lewandowsky et al. (2012)
C5	Pre-exposure warning		
C6	Information presentation	Situational Awareness	Endsley (2018)
C7	Information confidence		
C8	Information integration		
C9	Goals and motivations		
C10	Promoting truth in group setting		

Endsley (2018) reviewed state-of-the-art of fact-checking and provided framework to develop human-centered interventions. She claimed that fake news manipulates our cognitive consciousness by degrading situation awareness (SA), thus its solution requires 'cognitive engineering'. By reviewing relevant literature, she proposed several key factors which could be applied to design solutions that overcome our cognitive biases. Her framework includes factors like enhancing information presentation visually and integrating them, addressing people's inherent motivations and mental models, overcoming confirmation bias etc.

Information presentation was the most important criteria that could make people more objective about world events (Endsley, 2018). Urgent and repeated presentation of facts could mitigate backfire effect and reinforcement of false information. Misinformation should not be mentioned repeatedly in the correction. Simplicity was another important factor that makes information processing easier, which in this case would be presenting one or two fact-checking arguments concretely instead of lengthy ones (Endsley, 2018). Information framing is useful to minimize backfire effects of fact-checking. Choosing right words or right narrative can increase the likelihood of accepting the facts, for example replacing 'carbon offset' with 'carbon tax' (Endsley, 2018). Text-based warning of false information was also found effective when presented with correct information.

Endsley (2018) proposed information confidence as another major factor in developing fact-checking services. Information confidence means encouraging readers to be skeptic about the source of news. People are unlikely to believe false information if it comes from totally unknown sources. Tools that help people to manage information from different sources various reliability levels could be useful here (Endsley, 2018; Lewandowsky et al., 2012).

Fake news often comes in disorganized manner with inconsistent information that is difficult to process. Thus, their correction should be presented in a visual manner that makes information more interactive (Endsley, 2018). Graphical presentations could provide clarity on the topic and help readers to develop a more objective mental models. Endsley (2018) gave an example of news about climate change, where graphical representation can significantly increase its believability. However, graphs and visualizations should be presented with relevant context that is understandable by broad number of readers.

It's very difficult to change people's inherent motivations, and people with defensive mindsets might reject information that contradicts with their worldview. Based on Lewandowsky's (2012) findings, Endsley (2018) suggested to direct fact-checking efforts towards people who are less biased or less entrenched. Alternative narrative can be used to portray information in new lights or expose logical flaws in misinformation. While presented in an interactive way, this approach positively improved people's mental models (Endsley, 2018).

Overcoming confirmation bias and social forces are two major challenges in fact-checking. Cognitive bias could be minimized by encouraging people to form truthful opinions, especially if those come from their social group leaders (Endsley, 2018). Other researchers proposed interventions that encourage deeper thinking instead of directional reasoning. Changing opinions in group settings is still an under-developed research topic, but Endsley (2018) proposed to frame messages in a way that is consistent with social groups' values, even if that requires accepting them in some controversial issues. Some mitigation strategies taken by social media companies such as YouTube, Facebook was

found effective such as providing news context, which could be incorporated into other fact-checking tools (Endsley, 2018).

2.4 Persuasive technologies

Fogg's (1998) seminal work on human interaction with technology found persuasion as a major factor behind attitude and behavior change. Persuasion was defined as an attempt to influence human attitudes or behaviors or both, without using deceptive methods (Fogg, 1998). Interactive technologies designed for such behavior change are called persuasive technology. Systems built on this framework could be applied to wide range of domains such as education, business, health, professional activities, welfare (Fogg, 1998). Indeed, empirical findings proved that persuasive behavior models are highly effective in attitudinal and behavioral change (Hamari et al., 2014; Orji and Moffat, 2018). Systems based on these models have seen wide success in healthcare. Persuasive systems designed as interventions helped people to adopt healthy lifestyle by changing or reshaping unhealthy habits such as smoking, substance abuse, over-eating as well as improved mental condition for depressed patients (Orji and Moffat, 2018).

Persuasive technologies have been used to encourage certain types of behavior among users. Forget et al. (2008) used persuasion to influence users into creating stronger passwords. They found that motivating users to increase password security is more efficient than overwhelming them with manual instructions. Wunsch et al. (2015) designed persuasive strategies to encourage people into biking, which returned promising results. Oeldorf-Hirsch and Sundar (2010) found gratification techniques that motivate people to share photo online. Many persuasive systems have been used to influence people's behavior towards sustainable living and energy saving (Shih and Jheng, 2017; Stribe and Larson, 2016). However, no fact-checking service has applied persuasive strategies to fight the problem of fake news so far.

2.4.1 Persuasive Systems Design (PSD) Model

While Fogg's (1998) persuasion theory provides meaningful understanding of persuasive technology, its conceptual nature makes it difficult to directly apply in development or evaluation of persuasive systems (Oinas-Kukkonen and Harjumaa, 2008). Persuasive Systems Design (PSD) is a comprehensive framework that can be used to design and evaluate persuasive software. The model provides a structured approach to persuasive systems that are contextualized and targeted for specific behavior change. It offers a set of postulates that describe the key characteristics of persuasive systems, persuasion context and design principles that can be mapped into system features (Oinas-Kukkonen and Harjumaa, 2009).

The first step in designing persuasive systems is acknowledging the seven postulates, which were derived from prior researches in cognition and psychology field. The first postulate describes IT as never neutral, and it always influences human behaviour in a way or another. Second postulate states that people like their worldview to be consistent and organized. Third postulate asserts that persuasion can follow direct or indirect route, so persuasion strategy should be designed according to individual's existing attitude. Fourth postulate states that persuasion should happen in a gradual process instead of monolithic effort. Fifth, sixth and the last postulates are about openness, unobstrusiveness and usefulness of the system (Oinas-Kukkonen and Harjumaa, 2009).

After addressing the postulates, persuasion context should be analysed. This is done by identifying the intent behind persuasion, determining persuasion event and recognizing

the persuasion strategies in use. Intent identifies who is the persuader and what type of attitude or behaviour change is targeted. As machines do not have intentions of their own, there are always human persuaders behind systems. Intent is implemented in systems through persuasive event. In this phase, context of use, user and technology is recognized to understand persuasion event. Persuasion strategy is identified in the last step, which reveals which route, direct or indirect, is followed for persuasion (Oinas-Kukkonen and Harjumaa, 2009).

While postulates and persuasion context prepares the ground for persuasion, they are not specific enough to be used as system requirement. Thus, Oinas-Kukkonen and Harjumaa (2009) provided a set of design principles that can be used to develop and evaluate functionalities of persuasive systems. These principles were divided into four categories: primary task support, dialogue support, system credibility support and social support. Primary task support, the first category, helps the users to accomplish a primary task by reducing complex steps or monitoring their performance. For example, heart rate monitors could encourage users to exercise more. Dialogue support features establish communication with users by providing feedback through rewards, appraisals, and suggestions. An example is fitness applications that gives virtual badges as rewards to the high-performing users. System credibility support features lean on credibility and trustworthiness to persuade users. For example, a system can show certificates from third-parties to improve its credibility. The last category is social support features, which leverage from social attitudes such as comparison, competition and facilitation. Example of this category is systems where a user can compare their performance with other users (Oinas-Kukkonen and Harjumaa, 2009).

2.4.2 PSD analysis of fact-checking services

To identify persuasive features in existing fact-checking services, an analysis was conducted using the PSD model. Services for review were selected from several literature such as Brandtzaeg et al., (2017) and Brandzaeg et al. (2018). Popular fact-checking services of different types were selected which represent that particular domain according to Table 1. The following table shows the analysis:

Table 3: Analysis of existing fact-checking services using PSD model

	Principles	Hoaxy	Snopes	PolitiFact	T.Trails	ClaimB	FB
Primary Task support	Reduction	X	X	X	X	X	X
	Tunnelling	X					
	Tailoring						
	Personalization						X
	Self-monitoring						
	Simulation			X			
	Rehearsal						
Dialogue Support	Praise						
	Rewards						
	Reminders						
	Suggestion						
	Similarity			X			
	Liking		X	X			X
	Social Role						
System Credibility Support	Trustworthiness	X	X	X	X	X	X
	Expertise	X	X	X	X	X	X
	Surface Credibility	X	X	X	X		X
	Real Wrld. Feel	X	X	X			
	Authority			X			
	3rd party	X	X	X	X		X
	Verifiability	X	X	X	X		X
Social Support	Social Learning						
	Social Comp.						X
	Norm.Influence						
	Scl. Facilitation	X			X		X
	Cooperation						
	Competition						
	Recognition						

From the table, we can see that system credibility was the only prevalent persuasive feature category. In this category, trustworthiness and expertise were most common features among all services. However, no service used any dialogue support feature, which made it the most under-utilized category. Lack of dialogue support means the experience in these services were not interactive at all.

All services applied at least one primary task support feature which is reduction. This means all of them tried to provide an answer about authenticity of news by checking the

facts by themselves. Self-monitoring was proven highly effective in healthcare, which was not applied by any service.

Presence of social support features were also minimal. Hoaxy, TwitterTrails and Facebook applied social facilitation by showing info (view, shares) about other users. Lack of social support features in Facebook was quite surprising as it's the largest social network in the world.

Among independent fact-checkers, only PolitiFact has a gamified mobile application. While the game helps with primary task support, it doesn't have any social or dialog support feature. But presence of the game indicates that some fact-checkers noticed the problem of user engagement, thus trying out interactive mediums like mobile apps.

2.5 Credibility and value paradoxes

Fact-checking services are developed as applications to be used by general internet users, which makes it part of the social web. Oinas-Kukkonen and Oinas-Kukkonen (2013) defined social web as the phase where internet got redirected towards people, replacing its traditional structure with human-centered applications. This shift gave birth to several paradoxes that users frequently face while using the social web. These are: privacy paradox, identity paradox, credibility paradox, friend paradox, filter paradox and value paradox. Among these, credibility and value paradoxes are relevant for the fact-checking system developed in this study.

Credibility paradox is associated with users' tendency to trust all content of the web, regardless of their source. Because web generates a massive amount of information, it's easy to run into facts that are made-up, inflated or opinionated. But people still consider web as trustworthy because of its openness and accessibility. Gaining and sustaining their trust is the key to achieve credibility, which is important to maintain reputation in the social web (Oinas-Kukkonen et al., 2013).

Value paradox is related to the paradoxical nature of web's promise. Although it provides ample opportunities to establish new social connections or create new businesses, their value diminishes because of the flat structure of the web. Unimportant information and activities could be perceived as important, and people could develop a false sense of identity with illusionary audience. It's also easy to contaminate the web environment with unnecessary information, which only grows with time without any implication for the real world. Thus, products and services launched in the web should have clear value proposition for the users (Oinas-Kukkonen et al., 2013).

3. Research Method

Design Science Research Method (DSRM) was used as main research approach to carry out this study. The reason for choosing DSRM was it fits the purpose of the study, which is to build an artifact that solves a particular problem (Hevner et al., 2004). In this chapter, theoretical background of design science is discussed as well as the implementation process for this study.

3.1 Design Science in Information Systems research

Information systems research consists of two broad paradigms: behavioral science and design science. Behavioral science paradigm has its origin in natural science, thus theory-building in nature. Researches from this paradigm develop and validate theories that try to understand or predict human and organizational phenomena in information systems setting. Design science, on the other hand, has its roots in engineering research and seeks to broaden the human and organizational capabilities by developing new and innovative artifacts. These paradigms together set the IS discipline at the intersection of people, organizations and technology. Thus, it is important to apply both paradigms complementarily to solve the complex problems of information systems and to make significant contribution to the discipline (Hevner et al., 2004).

Hevner et al. (2004) argued that behavioral science research often studies an artifact that has already been implemented in organizational setting, while design science tries to resolve identified problems of organizations by developing and evaluating the artifacts. These artifacts have been broadly defined as "constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices) and instantiations (implemented and prototype systems)". They may come in different forms, but are represented in a structured way so that they can be evaluated using quantitative or qualitative methods. Just like field experiments help behavioral scientists to understand organizational phenomena, design science researchers address the problem through the process of constructing artifacts (Hevner et al., 2004).

A conceptual framework was further developed by Hevner et al. (2004) by combining behavioral and design science. The framework helps IS researchers to understand, evaluate and execute design science research. It contains three main components: Environment, IS Research and Knowledge base. Environment is the space where people, organizations and technology comes together to address the business needs that is relevant to the research. Development or justification of theories or artifacts take place in IS Research. Knowledge base comprises of foundations and methodologies of previous IS research. Frameworks, theories and tools from prior research are foundations that use methodology as a guideline to evaluate or justify artifacts to achieve research rigor. Knowledge base continuously gets additions from research phase while the Environment gets applications to meet the business needs (Hevner et al., 2004).

As design science is by nature a problem solving process, building and applying an artifact to solve the problem gives designers deep insights, which can reveal effective ways to solve the problem. Hevner et al. (2004) provided seven guidelines to conduct such meaningful design science research in IS. They are following:

1. Design as an artifact: Outcome of design science research should produce a novel artifact in the form of construct, method, model or instantiation. The artifact should solve the problem in an innovative way.

2. Problem relevance: The designed artifacts should try to solve a problem faced by an organization or community. Importance and relevance of the problem should be considered instead of just validating existing theories or predicting phenomena.
3. Design evaluation: The artifact should be tested rigorously with an established evaluation method. Evaluation metrics and data collection methods should be well-defined. A proper evaluation should measure the artifact's quality, efficacy and utility.
4. Research contributions: Design science research should add value to the existing artifact, foundation or methodologies through tangible and concrete contribution.
5. Research rigor: Rigorous methods based on existing knowledge should be used in both design and evaluation process. Rigor should be also balanced with relevance.
6. Design as a search process: Design should be viewed as an iterative process where the solution is achieved through recurring evaluation and improvement.
7. Communication of research: Outcome of design science should be presented in an effective way so that both technical and non-technical audience can understand.

3.2 Design Science Research Methodology (DSRM)

Peffers et al. (2007) provided a comprehensive framework based on design science principles to conduct design science research. They emphasized the need of a commonly accepted framework in the domain. Named Design Science Research Methodology (DSRM), the framework consists of principles, practices and procedures to conduct design science research. The research should also meet three objectives. It should have consistency with existing literature, should provide a nominal process for research and should provide a mental model to present and evaluate the research.

Peffers et al. (2007) also presented a process model consisting six activities to conduct design science research. The model is a synthesis of previous research. Six activities are problem identification and motivation, objectives, of the solution, design and development, demonstration, evaluation and communication. Figure 1 demonstrates a visual representation of the model:

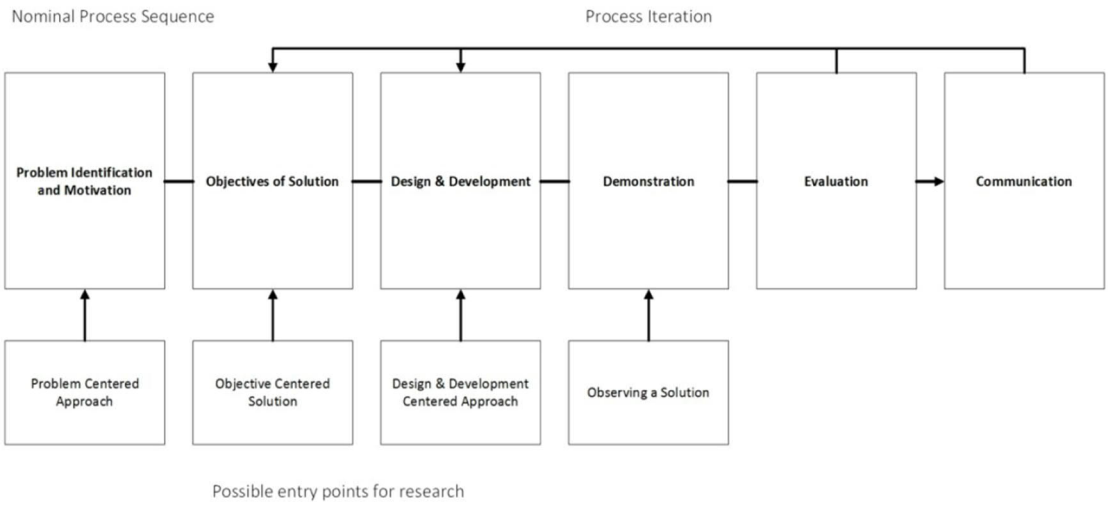

Figure 1: Design science nominal process model (Peffers et al. 2007)

The first activity is problem identification and motivation, which defines the specific problem and the rationale for a particular solution. Main resources of this activity are to know the state of the problem and importance of the solution. The second activity is about defining objectives for a solution, which identifies scope of the research and set objectives. These Objectives can be either qualitative or quantitative. The third activity is design and development, which involves identifying functionality and architecture of the artifact as well as creating the artifact itself. It could be a construct, a model, a method or an instantiation. The fourth activity is regarding demonstration of the artifact. With sound knowledge about usability of the artifact, researcher can experiment or simulate it to show its efficacy. In fifth activity, evaluation, the artifact is measured against the objective of the study to see how well it solves the problem in research. Based on the type of the solution, evaluation method can be quantitative or qualitative. Researches can also iterate back to third activity if unsatisfied with evaluation results. Communication, the final activity, involves presenting the problem, the solution and research results to the appropriate audiences. Researchers from the domain and practicing professionals could be audiences. Researcher can then use the empirical research structure to structure the paper and publish it in scholarly journals (Peffers et al., 2007).

The DSRM process is non-linear, which means it has different on-boarding points. Researchers do not have to conduct the research in the exact sequential order of the process. They can start from a certain activity and move forward. A problem-centered approach starting from the first activity would be suitable if the research stems from observation of the problem or suggested direction from previous literature. An object-centered approach can be taken if the research is motivated by industry or consultants. If the artifact already exists and solve some other problem than the problem of research, the process can start from design and development. It can also start from observing an existing artifact. In this case, demonstration activity can be the entry point and research can go backwards to increase rigor and validity (Peffers et al., 2007).

3.3 Implementation of DSRM

DSRM has been implemented to conduct this study. Problem-centered approach was followed due to the nature of the study, and it followed the sequential activities of the process. Basis of the research was the need of a behavioral intervention to combat fake news. Figure 2 illustrates the problem-centered approach applied in the study followed by an overview of the activities:

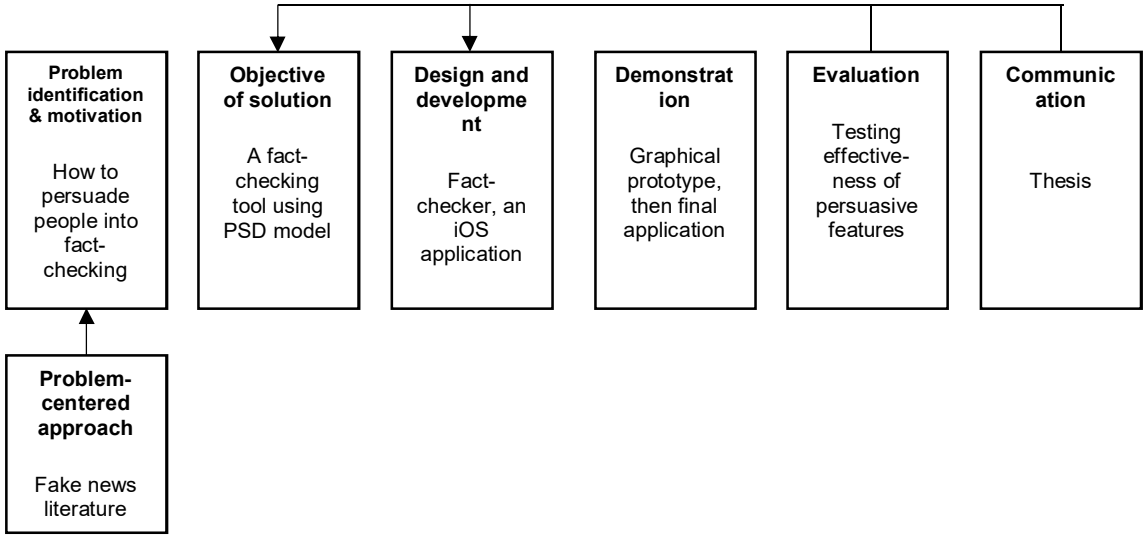

Figure 2: DSRM process implemented in this research

Problem identification and motivation: The problem was identified from state-of-the-art of fake news literature. Many fact-checking tools have been developed to deal with the problem, but most of them overlooked the cognitive aspects of fake news. PSD model proposed by Oinas-kukkonen and Harjumaa (2009) could be applied here to develop more interactive systems that persuade users into authentication of news.

Objectives of Solution: The objective was to develop a fact-checking mobile application that would persuade people to frequently fact-check news. Major difference with existing fact-checking tools is that our tool emphasizes on attitudinal change towards fake news, not on the technical know-how of fact-checking.

Design and Development: The artifact was developed by following standard software engineering process. Features of the software were rigorously identified through persuasive principles and relevant fake news literature. Requirement specifications were identified through use case diagrams and architecture developed using modern application development technologies. React Native, an open-source cross-platform framework based on JavaScript, was used as the main technology.

Demonstration: The artifact was first developed as a graphical prototype to understand user interactions. It was then developed into a full-fledged mobile application which could run in iOS devices. During development it was demonstrated to several users to redefine user experience.

Evaluation: The artifact was evaluated to measure effectiveness of persuasive features in fact-checking domain. Users were heuristically interviewed through a semi-structured interview and structured questionnaire. Results were be analysed and documented.

Communication: Result of the study was be communicated through this thesis. The entire research process including literature review, theoretical framework, research method, implementation process, evaluation, results and discussions was documented in the thesis, which would be available online for reading.

4. Software Artifact

A software artifact called 'Fact-checker' was developed for this study to investigate persuasion in fact-checking environment. It's an iOS application that tries to persuade users into fact-checking through behavioral interventions. Features of the application were identified through PSD model and cognitive psychology literature. Design and implementation process of the artifact is discussed in this chapter.

4.1 PSD model analysis

Persuasive System Design (PSD) model consists of three phases: persuasion postulates, persuasion context and software design principles. As basis of the design process, the seven postulates have been adopted which helps us to understand overall scope of the system. We acknowledge that the system will affect users' behavior in a way or another as information technology cannot be neutral. Complying with second postulate, the system will not directly challenge users' existing worldviews or opinions. It will take a more nuanced approach instead. Both direct and indirect features of persuasion will be followed and they will try to persuade users incrementally instead of monolithic efforts. User experience of will be intuitive and easy-to-use, and users will be clearly notified about the purpose of the system.

4.1.2 Persuasion context

Persuasion context analysis is required to recognize cognitive consistency of a system. A careful analysis of context could provide a consistent user experience. The analysis is done in three steps: the intent, the event and the strategy.

Intent

Identifying persuader and the type of intended change is the way to recognize the intent of the system (Oinas-Kukkonen and Harjumaa, 2009). Persuader of the system was the designer of the application, who is a student researcher in the Information Processing Science department of University of Oulu. It's designed to be autogenous, which means it could be used by the users without help of the experts. Alternatively, persuaders could be creators of existing fact-checking services who want to change their users' behavior towards fact-checking.

Intention of the system was to develop fact-checking behavior in users. It would encourage users to verify truthfulness of news and engage users in the fact-checking process. According to the discussion in section 2.2 and 2.3, such interventions are required to overcome the cognitive biases of fake news.

Another intent of the system was from research perspective, which was to study whether persuasive features could change users' behavior or attitude towards fact-checking.

Event

User context: The application was intended for individuals who frequently consume online news and share them in social media. In a way, user could be anyone who uses internet. But this broad scope also makes it difficult to tailor the application for certain groups. Thus, the application focused on users who were already conscious about the problem fake news and wanted to fact-check news before sharing, but found current fact-

checking services difficult to use. In summary, we assume that the user already has some inherent motivation to verify news before sharing.

Use context: Use case of the application was whenever users come across an online news they are skeptic about. They would visit the application to check truthfulness of the news. Ideally, they would not be motivated enough to check authenticity even though they are skeptic, because such verification requires additional cognitive effort as discussed in section 2.3. This is where the application comes in. It would have persuasive features such as virtual rewards, points system and self-monitoring, which are expected to motivate users into fact-checking. Users might even frequently come back to the application to casually check latest false news roaming around in internet.

Technology context: The application was developed for iOS platform and it could be published in the App Store. Most users come across fake news through mobile devices, so a smartphone application would make the system available to most iPhone users. React Native, an open-source framework of JavaScript, was used for development.

The system itself did not have any fact-checking mechanism, so the news were fetched from a third-party APIs provided by the Google's fact-checking API. It should be noted that emphasize of the system is on persuasive features, not on the process of fact-checking.

Strategy

The system utilized both direct and indirect persuasive features. Source-checking and reward system were the most visible and direct persuasive feature, along with self-monitoring dashboard. Indirect cues involved persuasive messages in form of notifications and reminders.

4.1.3 Persuasive software features

Software features of Fact-checker were identified using persuasive design principles and cognitive constructs suggested in previous literature (Table 2). The focus was on the persuasive features that are absent from existing fact-checking services (Table 3). iOS platform from Apple was selected as the implementation medium, as high-modality interactive features could be easily designed and deployed in the platform.

An overview of major features is following:

Newsfeed

Newsfeed displays a real-time feed of potentially fake news that is roaming around in the internet. These news are fetched from another third party fact-checking service through APIs, which in this case was from PolitiFact, one of most popular non-profit fact-checking services. This is also the home page of the application.

The feed helps the users with the primary task support, which makes the task of fact-checking simple and efficient. Users don't need to find sources of a news by themselves as all content are clearly presented with sources and facts, which increases modality of the interface as suggested by Oh and Sundar (2015). Such integrated presentation of news has been suggested by Endsley (2018) as a mitigation technique of fake news.

Source-checking and rating

This is the most prominent persuasive feature of the application. Users can only rate news after they check source behind a news. If users click on the rating button before checking sources, they would be prompted to check sources first. Each time users rate news after checking sources, they would receive virtual praise and badges. This feature is core intervention mechanism in the application, backed my cognitive constructs proposed by Oh and Sundar (2015) and Endsley (2018) as shown in Table 2. It provides cognitive cues to the user to spend mental energy for fact-checking.

The app also shows how many people have viewed a particular claim, and how many of them rated it as true or false. If majority of the users mark a news as right or wrong, this would influence others to believe the fact even if that goes against their existing belief. This feature was conceptualized from Lewandowsky et al.'s (2012) suggestion on skepticism about news source, pre-exposure warning and Endsley's (2018) suggestion on promoting facts in group settings. It was implemented in the system using normative influence feature, a social support category persuasive principle.

Praise and rewards

A virtual reward system is integrated in the application. Users will get rewards each time they check authenticity of a news before sharing and confirm that they have checked the sources. They get extra points for sharing factual news. When they gain a certain amount of points they receive virtual rewards such as Inspector, Citizen Journalist.

Several persuasive principles under dialog support category are implemented through this feature, such as praise and rewards. As seen from Table 3, this category is significantly missing in current fact-checking services. Endsley (2018) proposed goals and motivation as major factor to subconsciously engage people into information acceptance. Virtual rewards can persuade people to frequently verify news and help them to accept conflicting information.

User statistics

Users can check statistics of their fact-checking behavior. They can measure how many news they fact-check each week or month, just like users can measure daily calorie intake in health applications.

This feature is implemented according to the self-monitoring principle under primary task support category. This is one of the most effective persuasive features in domains like healthcare. But almost no fact-checking service currently provides such usage statistics. The feature is also supported by Endsley's (2018) suggestion on information integration, where she asked to present information in a graphical way to reduce misinformation.

Notifications

The system often send notifications to users alerting them about potential fake news. However, it only suggests to check sources of the news instead of directly confronting them to avoid backfire effect. This falls under reminder type of persuasive feature. Endsley (2018) proposed this type of intervention under information presentation category and Lewandowsky et al. (2012) proposed this as pre-exposure warning to fake news.

User profile

Users can sign up for the application and have personalized profile pages. The page will contain basic information about the user. Although this is not primarily a persuasive feature, such profiling could give users a tailored user experience. Most fact-checking services mentioned in section 2.2 do not allow users to create personal profiles.

Table 4: Implemented features of the system

Category	Principle	Implementation	Cognitive construct
Primary task support	Reduction	The application reduces the complex task of fact-checking by providing a real-time newsfeed	C1, C8
	Self-monitoring	Users can monitor how frequently they check truthfulness of news	C6
	Tunnelling	Users can only rate a news after checking sources, which is expected to form fact-checking habit	C2, C7
Dialogue support	Praise	System will praise users with motivating words when they check source of a news	C9
	Rewards	Users will get virtual badges when they earn certain points and reach different levels	C9
	Reminders	Notification will be sent to users when a potentially fake news gains popularity	C5, C6
System credibility	Verifiability	Each news will contain original source from where the news has been fetched	C7
Social support	Normative influence	Shows who already viewed the news and rated as true or false	C4, C5, C10

4.2 Software development process

Standard software development process was followed to develop the system. Requirements of the system were identified first, which helped to decide the software architecture. A prototype of the system was developed before final implementation. They are elaborated in the following sections.

4.2.1 Requirement specification

The first step of development process is requirement specification, where functional and non-functional requirements of the artifact are identified. Stakeholders and their needs are identified in this step, as well as documenting the requirements for further design, analysis and implementation (Nuseibah and Easterbrook, 2000).

Stakeholders of the application are users who want to check authenticity of news through smartphones. Another stakeholder could be owners of existing fact-checking tools who want their users to develop fact-checking behavior. As most features of the system are automated, system owners' do not have much responsibility other than providing real-time, updated API that fetches popular claims. Users are required to sign up for the system and create a user profile with some personalized information name, email. Main use cases for the user are to check sources of a news and mark them as true or false. Another use case is to visualize their performance of fact-checking and compare it with other users. If they are looking for a particular claim, they can search for it through search feature. End-users are not allowed to add, modify or delete any content as they will be fetched from third-party API.

A high-level list of key requirements of the system is following:

Table 5: Functional requirements of Fact-checker

No.	**Title**	**Description**
Req-1	Sign-in	Users must be able to sign-up in the system and sign-in
Req-2	Browsing	Users must be able to browse recent claims/ news
Req-3	Source-checking	Users must be able to check sources of a claim and authenticate it based on available facts. They cannot rate the news before checking facts
Req-4	Reward system	Users must receive virtual rewards/ badges when they earn certain points through regular fact-checking
Req-5	Self-monitoring	Users must be able to monitor their statistics about fact-checking and must be able to compare that with other users
Req-6	Notifications	System should send user frequent notifications about recent online claims

4.2.2 Software architecture

The bridge between requirements and implementation phase is software architecture (Garlan, 2014). It's a high-level representation of the abstraction that leads to design and implementation of software (Garlan and Shaw, 1993). Complex systems like software consist of different abstraction and operation levels with individual architectures. Software architecture illustrates the configuration of these architectural elements and how they interact with each other. Role of appropriate architecture becomes even more essential as the system scales and gets complex (Garland and Shaw, 1993).

MVC architecture has been traditionally the most popular architecture for this type of system. However, traditional MVC framework has challenges such as managing the application view, especially when it deals with different DOM manipulations. Because MVC follows two-way data binding, it re-renders the app even for small change of element, leading to heavy UI (Paul and Nalwaya, 2016).

React was developed by Facebook as a JavaScript framework. The goal of the framework was to make complex user interface building easier, where data changes dynamically. React provides the V part of the MVC (model-view-controller) architecture. Based on principles of React, Facebook released another framework for building native mobile applications, called React Native. It takes a hybrid approach towards development, which means the application is built using web technologies here, but it's rendered, executed and displayed as a native application. Both React and React Native uses Flux architecture, where data flow is unidirectional. Instead of two-way data bind like MVC, they use a concept called Virtual DOM. This allows the application to keep pages updated with minimal load and refreshing (React Native; Paul and Nalwaya, 2016).

The main advantage of using React Native is that source code can be compiled for both iOS and Android platform. An abstraction layer called the 'bridge' enables Native to render APIs in Objective-C (iOS) or Java (Android). Because of this time-efficient, lightweight architecture and UI-centeredness, React Native was used as the main technology to build the system.

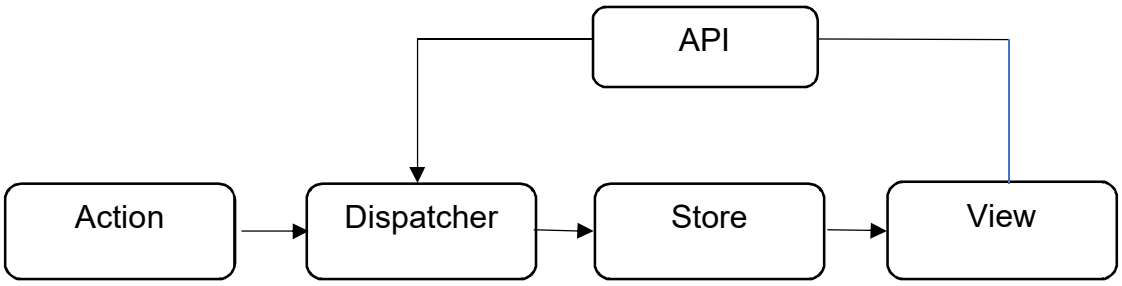

Figure 3: Flux application architecture for React Native (Flux, 2014)

Based on these analyses and requirement specification, following architecture was drawn for the system. The general architecture addresses the functional requirements as well as persuasive features of the application. Main content provider of the system is a third-party fake news/ online claims database. Claims will be fetched in JSON format through REST APIs, which will be parsed and displayed in the newsfeed. All other features of the system are based on these parsed content.

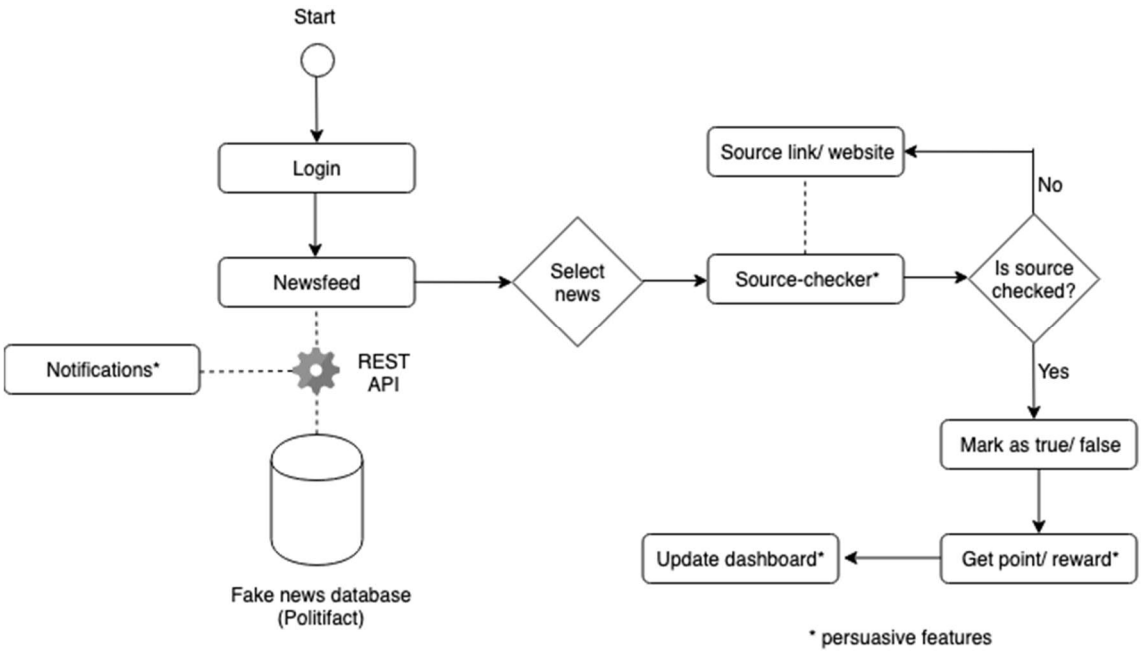

Figure 4: A general architectural view of the application

4.3 Implementation

Implementation of the system was conducted in two steps. A high-fidelity prototype of the application was developed first using an online prototyping tool. Final implementation was conducted through React Native technology.

Prototype, in design science context, has been conceptualized as an artifact that approximates the functionalities of a product, system or service. It has been present throughout history and interwoven into many technological breakthroughs. However, different prototyping efforts require different strategies, so the end-goal should be kept in mind while designing it (Camburn et al., 2017). In this research context, prototyping was done to better understand the workflow of the application and its different functionalities.

A simple and intuitive design was followed to build the system, according to Apple's guidelines (Apple). Because the audience of the application are general internet users, look and feel was kept as minimal and unobstructive as possible. Main components of the application such as user profile, statistics and newsfeed were accessible by the bottom navigation bar.

The first designed feature was a login page. Users need to create profile to use the system, so a login page was the entry point to the system. Users could sign-in with their e-mail and password. In the future versions of the application, social login with Facebook, Twitter could be added. Users need to sign-in for the application only once. They will remained signed for future uses.

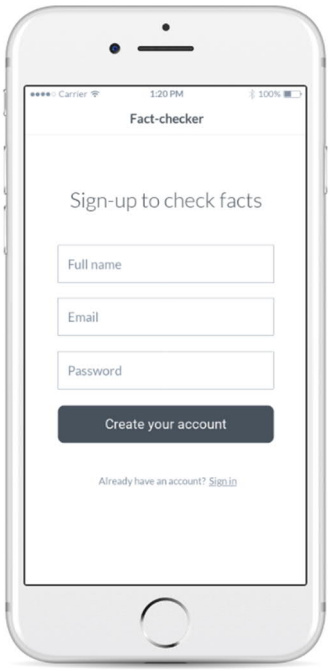

Figure 5: Sign-up page

Next step was to design the homepage of the application that serves the core purpose: fact-checking. Content of the page would be fetched from the third-party fact-checking API. Each content would be displayed as a news piece, where an online claim would be analysed with facts. Based on the analysis, news could be marked as true or false. Structure of the newsfeed page was similar to news applications where content is displayed in a list view.

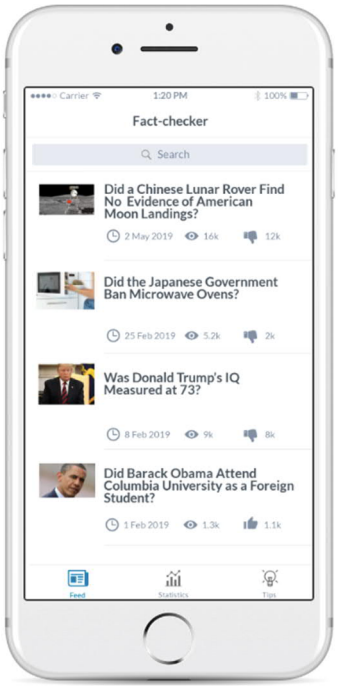

Figure 6: Newsfeed, homepage of the application

When users click on a content, they would be redirected to another page where fact-checking takes place. They mark the news as true or false in this page. But to mark the

news as true or false, source should be checked first by clicking the 'Check sources' button.

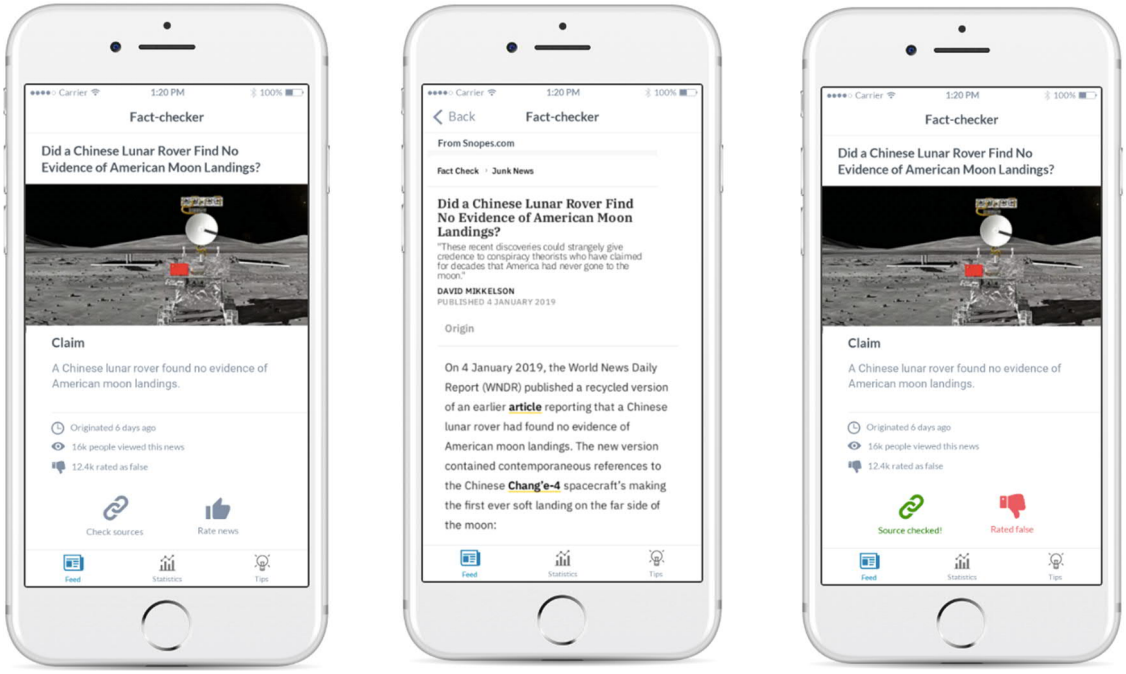

Figure 7: Source-checking of claims

A small snippet showed details about the content, such as when it was published, how many people have read it and how many have rated the claim as true or false. Persuasive features according to Table 3 are applied here to encourage users to check sources. If users mark the news against majority, a warning would be shown to check sources again. For each source-checking, users will receive rewards. After earning certain points, they would be awarded virtual rewards.

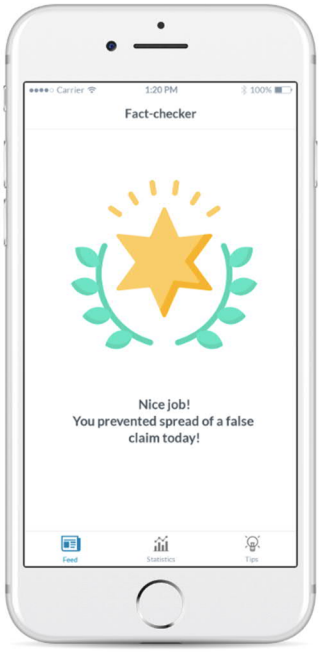

Figure 8: Virtual reward

In the statistics page, users could monitor their performance of fact-checking. They could check how many facts they are checking each week or month. This page implemented the self-monitoring persuasive principle.

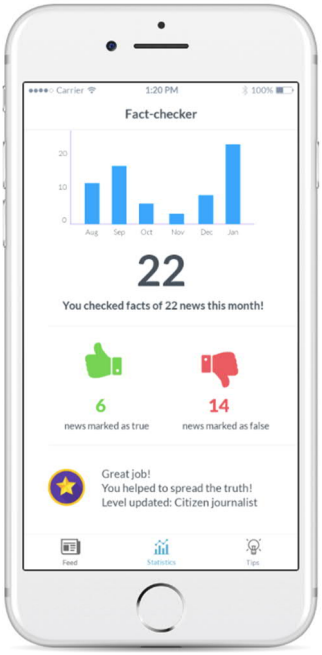

Figure 9: Self-monitoring dashboard

The first step of application development was setting up the environment. This was done with done through Homebrew, a package manager for macOS. Node package manager, a popular supplement of node.js, was used to install React Native CLI. Xcode IDE was required to run iOS applications, which was pre-installed in the computer.

React Native was installed with a simple npm command, followed by project creation.

```
npm install -g react-native-cli

react-native init FactChecker
```

The first step of the implementation was fetching the content from the third-party API. API of Politifact, a popular third-party fact-checking service, was used to fetch recent online claims. The JSON data was parsed with following code in app.js folder:

```
import React, { Component } from 'react';
import {
  AppRegistry, StyleSheet, Text, View, ActivityIndicator,
ListView, Image,} from 'react-native';

export default class MyFirstApp extends Component {
  static navigationOptions = {
    title: 'politifact',
  };

  constructor(){
    super();
```

```
      dataSource: new
ListView.DataSource({rowHasChanged:(r1,r2)=> r1!=r2}),
      link:
'https://s3.amazonaws.com/static.politifact.com/api'
  }
  componentDidMount () {
    return fetch(this.state.link)
        .then ( (response)=> response.json())
        .then( (responseJson) => {
            this.setState({
              isLoading: false,
              dataSource: responseJson;
              this.setState({
                dataSource:
this.state.dataSource.cloneWithRows(data)
              })
            })
        })
        .catch((error) => {
          console.log(error)
        });
  }
  render(){
    const { params } = this.props.navigation.state;
    return(
      <View style={styles.container}>
        <Text style={styles.pageName}>Newsfeed></Text>
        <Text> Category Name: {params.cat}</Text>

        <ListView
        dataSource={this.state.dataSource}
        renderRow={(rowData)=>
          <View><Text>{rowData.pro_name</Text></View>
        }
        />
      </View>
    );

   }
}
```

Newsfeed of the app was based on this JSON object. It returned title, image, description and source content of claims, from PolitiFact's database. When clicked on content of the newsfeed, users would go to the source-checking page, to check the original source from PolitiFact and then rate its truthfulness through the thumb-icon.

Core functionality of the app ended here. Rest of the features were implemented in the app as interactive mockups. Users could interact with them, but information would not be updated. Req-5 and req-6 were applied in the app with mockups.

5. Evaluation

Evaluation is one of the important activities in DSR. Hevner et al. (2004) emphasized on evaluation to test an artifact's quality, utility and efficiency. It helps researchers to identify contribution of the artifact to people, organization or discipline (Vaishnavi and Kuechler, 2004). It also helps to understand strength and weakness of an artifact as well as indicating future improvement areas (Casal et al., 1998). Venable et al. (2012) emphasized on rigor during evaluation, as rigor reveals whether the artifact serves its purpose. Without rigorous evaluation, one cannot prove if the artifact solves the problem or improves an existing solution.

The artifact developed in this study also went through the evaluation process. Objective of the evaluation, participants' selection, evaluation method and findings are discussed in this chapter.

5.1 Objective

The main objective of evaluation was to understand Fact-checker's effectiveness in influencing users' fact-checking behavior. The system was expected to encourage users into fact-checking and they should return to the system regularly to check authenticity of news. Because PSD model was applied to design the system features, evaluation would reveal the model's effectiveness in fact-checking context.

If the system was found persuasive, another objective was to measure which persuasive features were most effective in this context. Also, it was expected that evaluation would help the researcher to identify strength and weakness of the system as well as receiving suggestions for future improvement. It should be noted that evaluating technical aspects was out of scope for this thesis.

5.2 Participants

The first step in the evaluation process was selecting the participants for experimentation. Recommendations from prior scholarly work of Lewandowsky et al. (2012), Endsley (2018) and Khan et al. (2019) were followed to select participants who would be benefitted from the system most. Previous findings strongly indicated that fact-checking tools are most effective for users who already have some familiarity with the notion of fact-checking. An effort was made so select such participants who are aware of the fake news phenomena, and would give useful insights about the artifact. Another major criteria for selection was participants' familiarity with ongoing news trends, as most fake news or online claims usually stem from recent trends. Content displayed in the newsfeed were also based on recent hoaxes and claims, so participants were expected have some basic awareness about ongoing news trends.

A semi-structured short questionnaire (Appendix A) was used to find appropriate participants according to the selection criteria. Structure of the questionnaire was adapted from Khan et al.'s (2019) work, who developed a survey to measure people's information sharing behavior. Questions were adopted to fit the scope of this research. There were total six questions, among which four were based on 5-point Likert scale, and two were selection/open questions. Open questions were added to cross-check answers with Likert scale responses. At the end of the questionnaire participants were asked if they would be interested to participate in the main experiment.

The questionnaire was printed on paper. During a working day, it was randomly distributed to 30 students of University of Oulu Linnanmaa campus. Students were asked to fill-up the questionnaire with pen or pencil. They were notified that the survey was being conducted for a Master's thesis. 27 participants filled up the full questionnaire. Following table shows their responses. Mean is shown for first four responses. Answers of last two questions were qualitatively analysed to identify most relevant participants.

Table 6: Responses for the selection criteria questionnaire

No.	Question	Type	Mean (1-5)
FC1	Frequency of coming across fake news	Likert	4.13
FC2	Frequency of checking facts	Likert	3.90
FC3	Familiarity with fact-checking services	Likert	3.12
NT1	Familiarity with latest news trends	Likert	3.33
FC4	Method for verifying/ authenticating news	Select	n/a
NT2	Topic of a fake news trend	Open	n/a

Participants who had minimum score of 2 and were familiar with at least one recent popular fake news topic were asked for the main evaluation. Total 8 people met this criteria, and they were invited to participate in the interview. They were briefly explained the goal of the research and how their contribution would help to fight the problem of fake news. Later on 4 more participants were added to increase variation within the sample. These 4 persons scored average on Likert questions but did not recognize any significant fake news trend. In total, 12 participants agreed to be interviewed for the final round. For the interviewees who were not available immediately, a mutual time was fixed. Incentives were provided for their participation.

Table 7: Distribution of participants' age, gender and selection questionnaire score

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Age	34	25	28	26	27	31	25	25	23	25	27	24
Gender	M	M	M	M	M	M	F	M	M	M	M	F
Mean	2.00	3.25	3.75	4.25	3.75	2.25	3.75	3.25	3.50	3.00	2.75	4.50

In the final participants' pool, distribution of ages were between 24 and 34 (mean=26.68). Distribution of gender was 83.33% male and 16.67% female. Among the participants, majority were master's students (n=10), two was doctoral researcher. None of the participants were native English speakers, but all were fluent in English.

5.3 Method

The evaluation method consisted of three different experiments: usability testing, desirability testing and persuasive effectiveness testing. They were conducted through semi-structured interviews and structured questionnaires. Overall efficiency and effectiveness of the system were measured with usability testing, while desirability addressed user satisfaction and persuasive effectiveness addressed efficacy of persuasive features and intention to continue using them.

Usability testing

Usability has been defined as concept where a product could be used by particular users to accomplish a set of specific goals. Goals should be achieved with efficiency, effectiveness and satisfaction for the specific context in use (Jokela et al., 2003). To measure such efficiency and effectiveness of a system, Barnum and Palmer (2010) proposed usability evaluation. However, usability evaluation may not address overall user satisfaction of a system. Intangible experiences like enjoyment, desirability and fun are also difficult to measure with usability testing (Benedek and Miner, 2002). But it's an opportunity to observe users' emotions and feelings towards a system, so usability testing should be conducted with an overarching purpose (Barnum and Palmer, 2010).

Fact-checker's usability was evaluated through semi-structured interviews. Questions asked in the interview were prepared in Goal Questions Metric (GQM) method. GQM is a popular evaluation method for software features (Caldiera et al., 1994). In this method, questions are prepared based on goal statements which correspond to different components of a software, and metric tries to characterize whether the goal is achieved. Based on GQM, questions asked in the semi-structured interviews are following (more details in Appendix B):

Table 8: Interview questions for usability testing based on GQM

Goal	- To test the core intervention mechanism of the system - To identify which persuasive features users find most interesting and engaging in the application - To test if the system encourages users into fact-checking
Question	- What do you think of the source-checking feature? - Which features in the app did you like most? Which features would help you to check facts more frequently? - Would the app encourage you to be more conscious about fact-checking? Would it affect your fact-checking behavior?
Metric	- Measured overall usability of the system from the point-of-view of potential users

Desirability testing

Benedek and Miner (2002) developed a method to test desirability of products in Microsoft, called The Desirability Toolkit. It tries to test abstract feelings such as 'fun' and 'desire' in usability setting. The method invokes users to provide richer responses and reduces the risk of false positives, unlike traditional methods like Likert scales

(Benedek and Miner, 2002). While the usability testing helped to understand efficiency and effectiveness of Fact-checker, desirability testing was used to measure users' satisfaction with the tool.

Desirability toolkit from Microsoft consisted of two experiments: the faces questionnaire and the product reaction cards. In the faces questionnaire, respondents use the product first and then rate them by looking at pictures of facial expressions. In the product reaction cards, respondents choose few words from a pool that corresponds to their feelings after using the product. The pool should have a balance of positive and negative words so that users can express their negative bias as well (Benedek and Miner, 2002).

For this study, product reaction cards method was used as Barnum and Palmer (2010) found that this technique gives researchers more insights about user satisfaction than other methods. This is due to the fact that users are able to strongly express their emotions with product reaction cards (Barnum and Palmer, 2010).

After using the Fact-checker, 118 adjectives from product reaction cards were shown to the participants (Appendix C). They were asked to choose 5 adjectives that represent their feelings after using the tool. Explanation for selecting those adjectives were asked as well.

Persuasive effectiveness testing

The last phase of evaluation was testing persuasive effectiveness of Fact-checker. Lehto and Oinas-Kukkonen (2015) found that perceived effectiveness of a system is a strong indicator of whether the system will be continued to be used.

Lehto and Oinas-kukkonen (2012) provided an evaluation approach for persuasive systems that target behavioral change. They proposed a survey that contains measurement factors to test key persuasive principles. These factors are: primary task support (PRIM), dialogue support (DIAL), perceived credibility (CRED), design aesthetics (DESA), perceived persuasiveness (PERS), unobstrusiveness (UNOB), intention to continue using the system (INTE) and usage of the system (USE). Each of these construct contains one or several items which serve as a measurement instrument. All instruments were derived from prior scholarly work.

For this experiment, PRIM, DIAL, CRED, DESA, PERS and USE were selected for measurement as they were most relevant with the scope of the thesis. Following Lehto and Oinas-Kukkonen's (2012) structure, measurement items for each of the factors were constructed. Number of questions for each construct were slightly modified to accommodate in fact-checking context and to fit the short timespan of the interviews. In the final questionnaire, PRIM, DIAL, CRED, DESA, PERS and USE had respectively two, three, two, one, two and one measurement items.

Table 9: Evaluation constructs and their measurement items

Code	Construct	Measurement item
PRIM	Primary task support	Fact-checker helps me to check authenticity of recent online claims Fact-checker helps me to monitor my fact-checking statistics
DIAL	Dialogue support	Fact-checker provides me relevant feedback Fact-checker provides incentives/ rewards to check authenticity of news regularly Fact-checker notifies me about potentially fake news
CRED	Perceived credibility	Fact-checker is credible Fact-checker shows claims from authentic sources
DESA	Design Aesthetics	Fact-checker is intuitive and easy to use
PERS	Perceived persuasiveness	Fact-checker could influence my fact-checking behavior Fact-checker encourage me to think about authenticity before sharing a news
USE	Usage	I want to keep using Fact-checker

Another questionnaire based on Likert scale was prepared based on these constructs. After usability and desirability testing, participants were asked to fill-up the questionnaire (Appendix D). Each question was on a scale of 1 to 5 (strongly disagree to strongly agree). The evaluation process was completed with this final round.

5.3 Data collection

Data was collected through qualitative interviews and quantitative questionnaires. Both were conducted face-to-face with the participants, in a single session. Each session lasted for about 40 minutes on average. 8 of the participants were interviewed one-to-one, while rest 4 were interviewed in a group. The group interview lasted for one and half hour. All interviews took place in Tellus area and central cafeteria of University of Oulu.

Each interview started by thanking participants for their time. First they were given an overview of fake news and fact-checking tools. Then they were asked to use the Fact-checker application from the researcher's phone or from their own phone if they had an iPhone. A default profile with pre-set credentials was used for sign-in. After signing in, Participants explored the application for couple minutes by themselves.

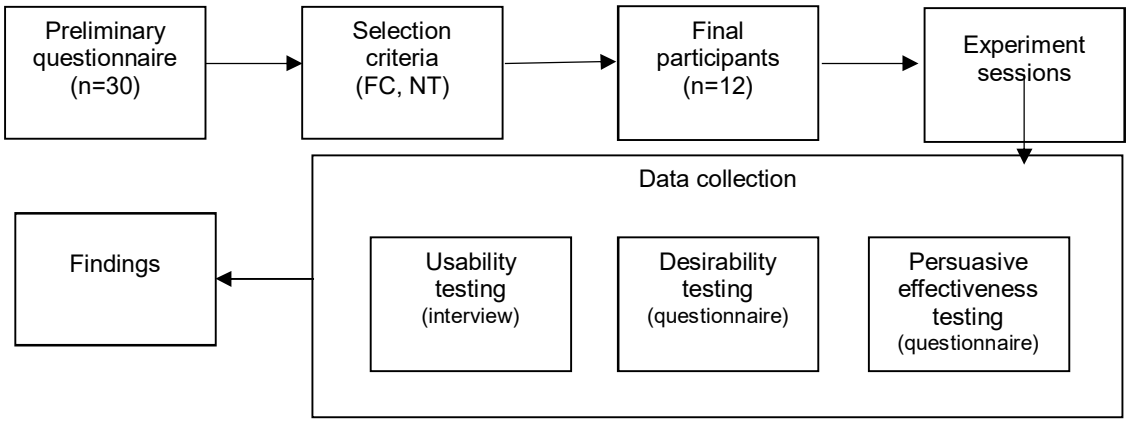

Figure 10: Conceptual model of evaluation process

After that, they were asked to perform a task, which was to check source of a news from the newsfeed and mark it as true/false. They could select any news from the feed or use the searchbox to find a recent claim. Most participants selected one of the default news pieces from the homepage. Then participants were asked to check the sources and mark the news as 'true' or 'false'. After checking source of the news, they marked it using the thumb button. Praising message was shown when they completed source-checking and rating, along with a virtual reward. During the process, they were asked to think aloud what they thought about different functions. After performing the main task, users were asked to view the self-monitoring and notifications features. These features were implemented as interactive mockups, so users could only interact with them but not change any information.

After using the app for about 10 minutes, main evaluation started. Usability testing was conducted first. Participants were asked questions according to table 6. Key points from their answers were noted. This phase took about 15-20 minutes. After that desirability test was conducted by showing them the adjectives on a paper. Users marked 5 adjectives they felt associated with the product. When marking each answer, they were asked to think loudly, which helped to understand why they were choosing that option. Desirability test took about 5 minutes to complete. The last experiment was persuasive effectiveness test, where users filled up a Likert scale questionnaire based on Table 7. This experiment took about 10-15 minutes per participant.

After completing all experiments, users were asked for their final remark on fake news and fact-checking in general. All interviews were completed in one week time.

5.3 Findings

Participants' responses for semi-structured interviews were collected in form of transcriptions. Responses to the questionnaire forms were collected and archived. After all materials were collected, they were qualitatively analysed by the researcher and compiled into a spreadsheet. Observations and personal notes from the researcher were also included. The overall process was guided by suggestions from Kaplan and Maxwell (2015) about coding data in qualitative research. After compilation, all data were synthesized in a readable format.

Usability testing

Usability testing was conducted through semi-structured interviews. All participants were able to complete the task of rating a news after checking sources. Three participants needed instruction about the rating, as they didn't understand they had to click the 'Check sources' button first before marking the news. When asked what they thought of the mechanism, 10 users were positive, while one mentioned that this was too complex and another mentioned this as time-consuming. The most positive comment from one user was:

"I think it's an excellent way of gatekeeping. You cannot just rate a news as true or false, you need check the source first. You're forcing people not to share stories without verifying—in a good way."

Another participant thought there was not enough incentive to check the source, and it's an unnecessary step. His comment was:

"It's an interesting feature, but also taking away my freedom. Why should I care about rating a news true or false if I don't care about the news at all?"

After the first question users were asked which features of the application they found most interesting and engaging. Interestingly, a unanimous answer here was the normative influence feature, which showed how many have rated that news as true or false. All participants said this would be the strongest indicator of whether a news is true or false. Four users said they would not bother to check the sources if majority of the users already marked the news. They would simply trust others' judgements. Although this was not the most prevalent persuasive feature of the application, most users noticed it right away and wanted to check it for other news pieces as well. One user wanted to check whether any of her friends rated the news, which unfortunately was not an implemented feature.

After social support, users mentioned self-monitoring dashboard as the most engaging feature. 8 out of 12 participants praised fact-checking statistics, while two said it wouldn't affect their fact-checking behavior. Two participant pointed out that current statistics was too simplistic. Instead of just showing self-monitoring performance, one of them recommended to add more meaningful and macro-information about the news:

"I want to see how many people shared this news, and how many actually fact-checked before sharing. This would motivate me to be part of that small number of fact-checkers."

Surprisingly, only two users recognized virtual reward, the major dialog support feature, as interesting. Most others said that rewards would be irrelevant in fact-checking context. Six pointed out that the virtual badges and praising texts would not incentivize them to check facts. Two users mentioned that the rewards shown in the prototype were too generic, and perhaps a better-designed gamification system would interest them. Among the two who were positive, one made an interesting comment that rewards could motivate him if those looked 'serious' and 'official'. He liked the use of word 'citizen journalist' in the praising message. The other dialog support feature, notifications, was mentioned by almost everyone as interesting. They wanted to be notified when a hoax starts gaining popularity in the internet. Two mentioned that they would be interested in notifications only if those are related to their country or community.

Verifiability was one of the persuasive features of the application, but it was not explicitly mentioned by any of the participants. Because the application domain was fact-checking,

it was expected that many users would express concern about verifiability. When asked, 6 out of 12 participants said they would like to know credibility of sources, even though they did not notice it while using the application. Answers from rest others showed they were less concerned about verifying source credibility. One interesting comment from one participant was:

"I would like to see the source of newsfeed content, from where you are getting the news. I understand it's an aggregator, but I don't know from which sources you are getting it."

After this comment, the API from which content were fetched was shown to the participant. He recognized the API as a credible source and asked to show the source within the system.

The last question in usability testing was what users thought about the application in general, and if it would encourage them into fact-checking. 7 out 12 subjects expressed strong interest in using the application, and said it would motivate them to regularly check facts. One subject mentioned that he would not be interested without personalized content. He further mentioned that the current newsfeed and persuasive features are too generic and irrelevant for his interest. Two users said they are not much interested in online news in general, so they would not be interested in fact-checking at all. Two users found the application somewhat useful, but time-consuming. One of these two suggested to add more social support features to keep it engaging.

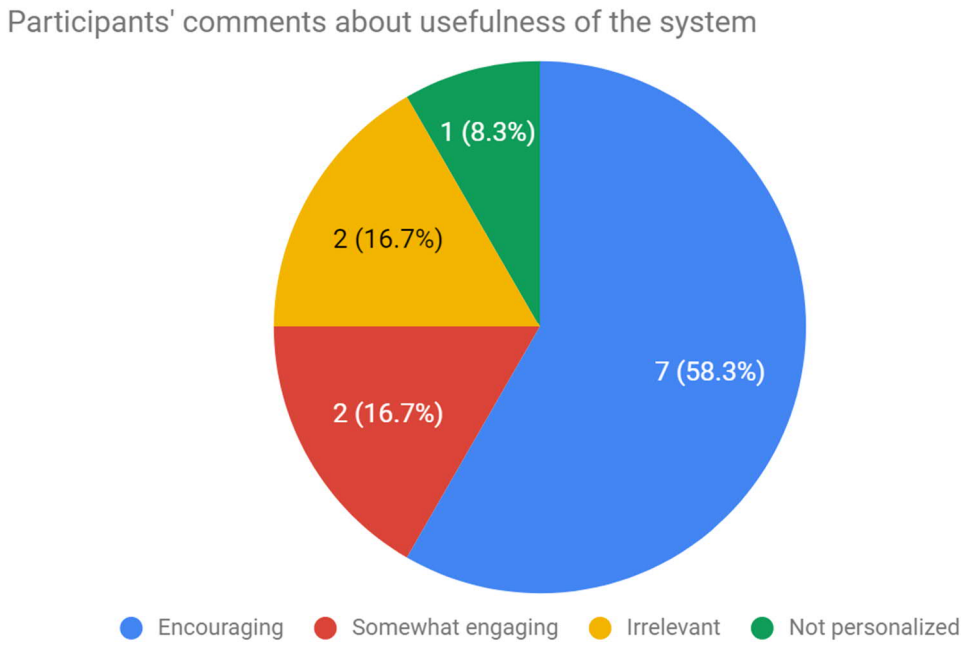

Figure 11: How participants perceived the system in general

Several functional and strategic usability issues came up during the discussion. A common question asked by most of the participants were regarding the source checking and rating mechanism. They asked if it would be possible to manipulate the ratings, for example, what if someone checked source but still rated the news from personal bias. Or, what if a group of people rated a news as false to manipulate public opinion. Another issue that came up several times was use context of the application. Many users mentioned that they would be more encouraged to use the system if the content was personalized and curated. Several users praised the simplicity of UI and intuitive user experience. One

viewed the whole concept as a news aggregator with fact-checking service built in. Another comment from several users were they would not the system as a standalone app, but as an integrated service with other popular news websites. Two users emphasized that the source-checking mechanism would be highly effective if integrated with existing news applications.

Desirability testing

After usability interviews, users were given a piece of paper with 118 adjectives. They marked a wide range of adjectives from the paper, which represents that the system was viewed from diverse perspectives. The following chart illustrates their selection:

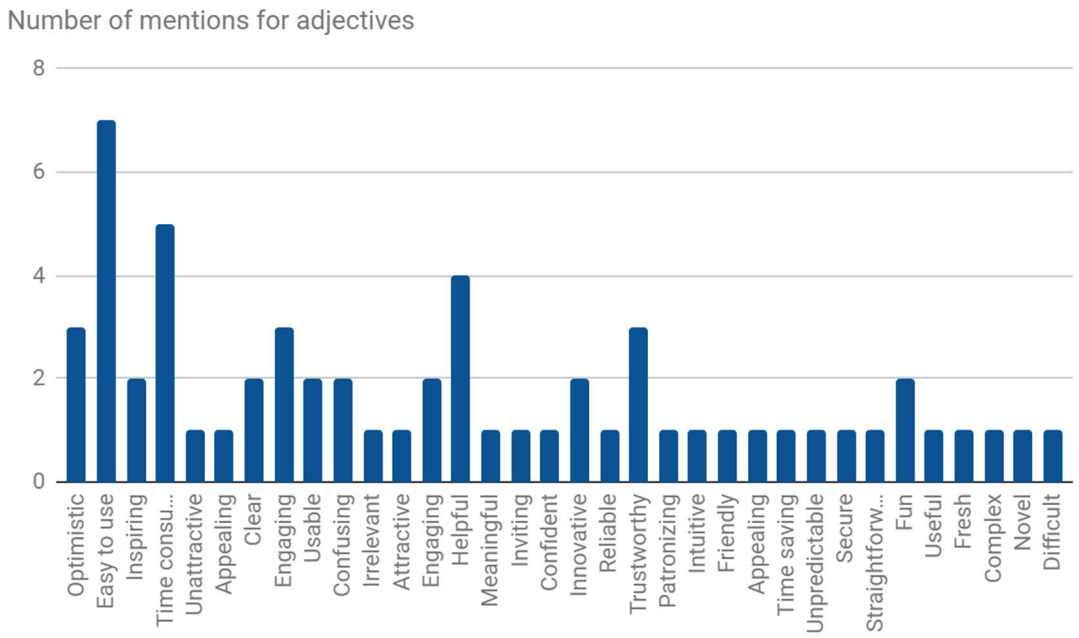

Figure 12: Distribution of adjective markings by participants

Out of 12 participants, 7 corresponded to the adjective easy to use, which means they were overall satisfied with the user experience. The second most chosen word was a negative adjective: time-consuming. Among other positive adjectives, helpful was chosen 4 times, engaging 3 times, trustworthy 3 times, optimistic 3 times and fun 2 times. Among negative adjectives, confusing was mentioned most after time-consuming. Difficult, unattractive, unpredictable and complex were selected one time each.

Participants were asked to think aloud while selecting the adjectives. The most common theme they mentioned was simple and intuitive user experience, which reflects the most popular adjective. Next emergent theme was user engagement by choosing keywords like appealing, engaging, helpful, fun etc. This means users found the overall system quite interactive and associated it with positive feelings. However, a common negative theme was complexity of source checking, which lead to a time-consuming experience.

Persuasive effectiveness testing

The questionnaire based on persuasive effectiveness testing was filled up by the participants as the last step of evaluation. Again, responses widely varied, but several patterns emerged after answers were compiled.

Table 10: Mean distribution of participants' responses

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
PRIM	3	3	2.5	4.5	3.5	4	2.5	2.5	4	2.5	3	4
DIAL	2.67	3	3	4.33	3	3.33	4	3	3	1.33	2.67	3
CRED	2.5	4	3	4.5	3.5	4	3.5	3	3.5	2	2.5	3
DESA	4	4	4	5	5	5	4	4	4	4	3	5
PERS	2	5	3	5	5	4	4	3	4	2	5	4
USE	2	3	3	5	5	3	5	4	5	3	5	5

Among all factors, DESA scored highest with mean of 4.25, followed by USE (mean=4.00) and PERS (mean=3.83). This result is quite consistent with the desirability test. Most participants agreed with the fact that the system provided a visually appealing experience along with smooth design aesthetics. Most users agreed or strongly agreed with this point. Six users expressed high interest to keep using the system by strongly agreeing with the USE question, while one disagreed and the rest moderately agreed. Perceived persuasiveness, an important construct, received above average score. Four users strongly agreed that the system would persuade them into fact-checking, while four moderately agreed, one disagreed and one was neutral.

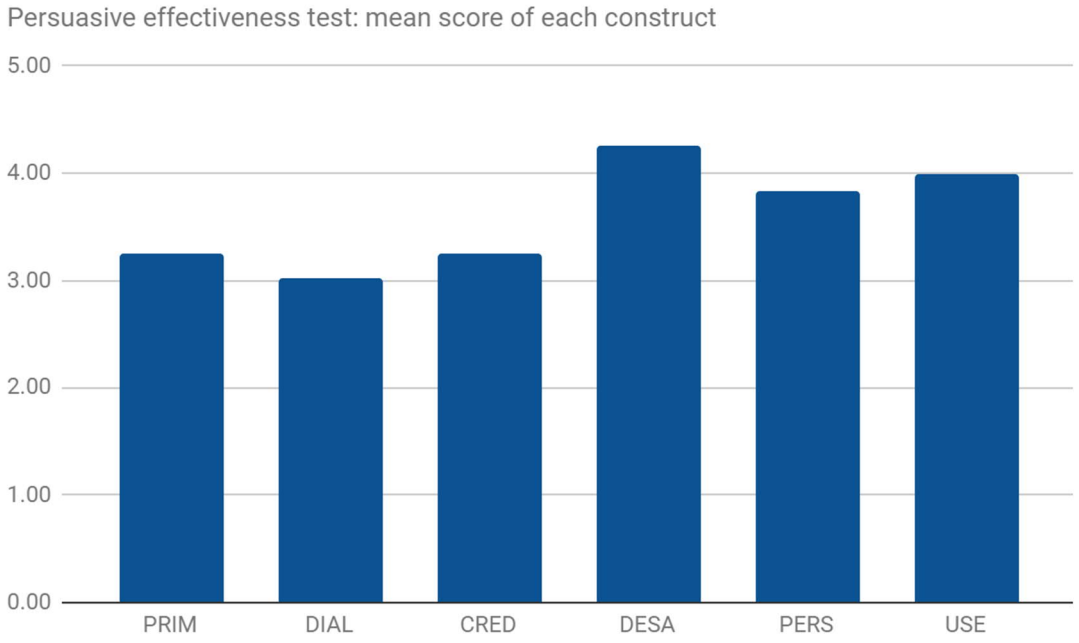

Figure 13: Mean score of different constructs in persuasive effectiveness test

PRIM category questions scored average with mean of 3.25. No participant strongly disagreed that the system helps them to check facts, but were neutral about its intensity. CRED factors scored same as PRIM (mean=3.25). Two participants were quite skeptic about credibility of sources, which reduced the mean here. The lowest mean was scored by DIAL factors (mean=3.03), where most users disagreed with the fact that the app

provides relevant feedback. No user rated ‘strongly agree’ to its questions, and two users mostly disagreed with all questions in this category.

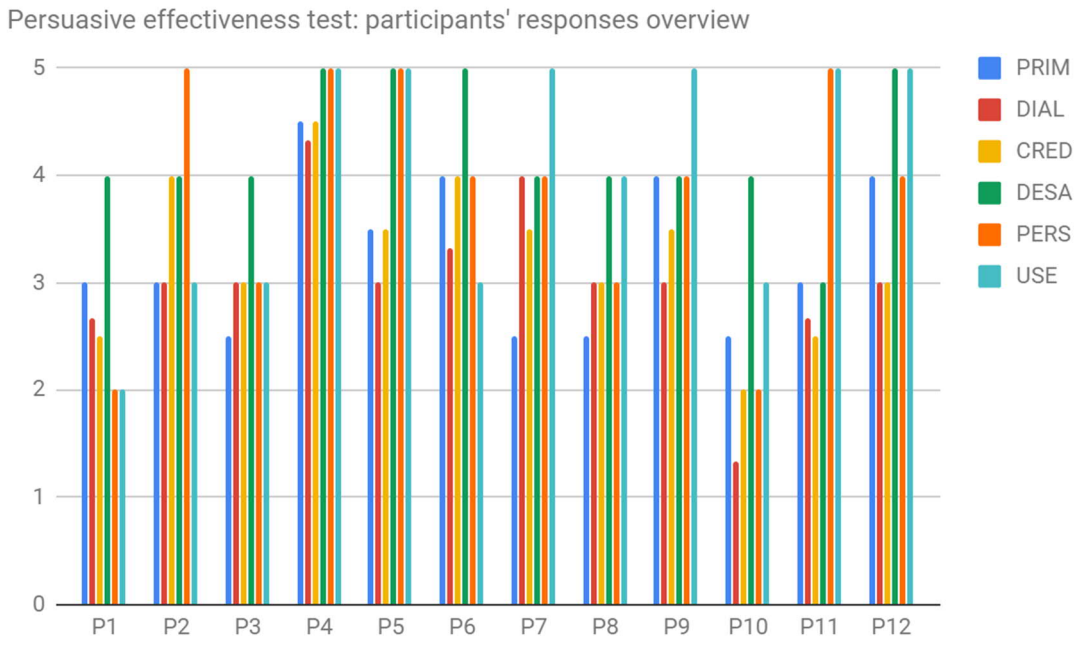

Figure 14: Overview of participants’ responses to the effectiveness questionnaire

Looking at the participants’ individual responses, it is evident that some participants had significantly better experience with the application than others. P4 was overall most positive about the persuasiveness of the application, followed by P5, P12 and P9. The system was least effective for P1, followed by P10 and P3. These participants responded negatively to most questions.

In section 5.3 we showed mean score of these participants’ preliminary questionnaire. By comparing means from that test with persuasive effectiveness mean for each participant, we get an interesting insight.

Table 11: Comparison of selection questionnaire mean with persuasive effectiveness

Participants	Selection questionnaire mean	Persuasive effectiveness mean
P1	2	2.70
P2	3.25	3.67
P3	3.75	3.08
P4	4.25	4.72
P5	3.75	4.17
P6	2.25	3.89
P7	3.75	3.83

P8	3.25	3.25
P9	3.5	3.92
P10	3.00	2.47
P11	2.75	3.53
P12	4.5	4.00

This table shows that participants who scored high in the preliminary questionnaire also tended to like the system, and participants who scored low in the questionnaire did not find the system much engaging. P4, who had the highest mean (4.25) from the selection round, also scored highest (4.72) in the persuasive effectiveness test. P1 had the lowest mean in both tests. From this observation, we could assume that users who are already conscious about fact-checking and news trends would find the application most useful and persuasive. This finding aligns with previous literature that it's easier to persuade people who are inherently less biased and less opinionated.

However, this particular finding from this thesis should be regarded as a general observation rather than an empirical finding, considering the biases in the experiments. Also, there were some participants who scored low in selection test, but scored high in the persuasive effectiveness test.

5.3 Summary

The evaluation process provided a holistic picture of the artifact. Goal of the evaluation was to understand Fact-checker's effectiveness in persuading people into fact-checking. From the qualitative and quantitative responses of 12 participants, the system was found quite persuasive comparing with existing fact-checking systems. Users found the application as an 'intervention' rather than just a content application. Ease of use, simplicity and straightforwardness were common themes that came up several times. On the downside, time-consuming nature of the application was criticized by several participants.

Measuring persuasive features revealed several interesting insights. Social support was the most prevalent persuasive feature, even though it was not strongly implemented in the application. Self-monitoring was perceived very positively as well, but improvements were suggested by the users. Source-checking feature, which was a combination of reduction and tunnelling principle, was highly promising. Not much interested was shown for the newsfeed and source credibility. Dialog support features performed poorly, and felt out-of-context for fact-checking systems. To summarize, the system was able to encourage a significant number of participants into fact-checking, but in a more complex and self-referential way than other persuasive systems.

6. Discussion

Evaluation of the artifact showed that persuasion could be an effective strategy to combat fake news. Instead of focusing on computational verification like most fact-checking tools, artifact of this thesis tried to test the cognitive nuances of news verification process. Recent literature reviews clearly pointed out lack of such humanistic approach. Thus, emphasize of this research was to find ways to closely engage users into fact-checking. Persuasive principles were used to design the software features, as systems built on persuasive technologies have been proven highly engaging and interactive, often leading to behavioral or attitudinal change. Fact-checker application was expected to form such regular fact-checking habit in users.

To achieve that goal, Fact-checker was developed based on two different research streams: persuasion and cognitive elements of fake news. Features of the application were derived from previous scholarly suggestions, particularly from Oh and Sundar (2015), Endsley (2018) and Lewandowsky et al.'s (2012) work. Endsley (2018) identified several cognitive avenues that are easily attacked by misinformation, and proposed a set of 'cognitive engineering' factors to build fact-checking solutions. Endsley's (2018) framework along with Oh and Sundar (2015) and Lewandowsky et al.'s (2012) suggestions were mapped with Oinas-Kukkonen and Harjumaa's (2009) persuasive systems design model to identify requirements of the system. A curated newsfeed, statistics dashboard and source-checking feature helped users with the primary task support, while dialog support features were implemented with virtual praise and rewards features. System credibility was implemented with verifiability feature, and social support was implemented with normative influence.

The thee-step evaluation process tested persuasiveness of the application as well as its overall usability and efficiency. While most users found the system quite persuasive, not all persuasive features actively engaged users. Dialog support features, for example, were found not much engaging, although persuasive systems from other domains successfully applied these features to influence users' behavior. Social support feature, interestingly, found unanimous applause from users. Not only users expressed strong opinion for this feature, they mentioned such social functions are absent from most existing fact-checking services, even though the social support feature implemented in the system was rather a simple one. According to the normative influence principle of this feature, the system showed how many people viewed the news and how many of them rated it. From the experiments, it was visible that this information could strongly influence users' fact-checking behavior.

This finding about social support features could be an important implication of this research. Cognitive biases and isolated echo chambers are the major challenges in fighting fake news, making it a complex socio-technical phenomena. Because of counter-intuitive nature of fake news, simply pointing out false information in a news is not enough. Numerous verification tools tried that, but failed to achieve sustainable results because of our psychological biases as discussed in section 2.3.

Social support persuasive features could be the nuanced strategy here to balance between fact-checking and cognitive biases. With social features like normative influence, users are not directly confronted with misinformation. Instead of forcing them to trust or distrust a news, social influence encourages them to stick with majority's judgement. As majority has already fact-checked the news as true or false, which is visible on the application interface, users are notified about its authenticity in a passive way. One could

refrain from sharing misinformation this way even if he/she is biased, to agree with the community's worldview. People's tendency to conform to social norms could be leveraged here, by making fact-checking a regular, visible and shared activity. In a word, normative social influence, the same phenomena that creates echo chambers with polarized opinions, could be used with indirect persuasion to create chambers where fact is prevalent. Experiments done with the Fact-checker artifact shows promise of such self-referential persuasive features. However, a major risk with this feature is it could be manipulated easily to form inflated public opinions. Further technical analysis is required to understand feasibility of this feature at scale.

Among other features, source-checking got significant attention from the users. The feature was a combination of reduction and tunnelling type persuasive principles. It helped users to check sources of news easily, but also enforced them to mark truthfulness of the news after checking sources. The gatekeeping nature of this feature was praised my most participants, but also added complexity to the system. Virtual points and rewards were provided in the system to motivate them into source-checking, which were not effective. They preferred a simpler mechanism to check and rate sources. But checking sources always mean there would be one more step before sharing a news, which means the notion of source-checking and rating is paradoxical in nature. But several participants suggested that a well-designed gamification strategy could explore the full potential of this feature.

Based on the experiment results and participants' comments, we could assume that Fact-checker would be most effective when integrated with other systems. Many users acknowledged the usefulness and novelty of the system, but also perceived it as an extra task or burden. From the point-of-view of users, fact-checking is a secondary activity, the first being collecting facts itself in form of online news. Unlike other persuasive systems, the goal of the user is not to check the facts, but to read factual news. This compliance attitude towards fact-checking makes it difficult to motivate people. But when these features are implemented in platforms like social media, they could be a powerful solution to the problem. For social media like Facebook and Twitter in particular, tunnelling, normative influence and self-monitoring features could be highly relevant in combating fake news. Computational and journalistic fact-checking tools could also integrate these features to increase persuasiveness of their systems.

During the experiments, it was evident that subjects' prior exposure to fact-checking had strong effect on their perception about the system. Participants who were familiar with recent news trends and aware of fake news epidemic understood the purpose the system instantly. Some of them gave deeper insights in the interviews and shared incidents where they were victims of fake news. On the other hand, participants who showed little interest with news trends in general perceived the system as unnecessary, and spent little effort in understanding the system. This observation aligns with the previous finding that fake news spreads from users' low cognitive effort. Thus, users with higher cognitive absorption should be the first target audience for such system.

As Fact-checker was developed for general online users, it was prone to credibility and value paradoxes as discussed in section 2.5. Credibility was expected to be the major paradox due to the nature of the application, and it was assumed that users would be concerned about credibility of the content as well as its sources. Surprisingly, most of the participants were not conscious about validity of the content. Since the goal of the application was checking credibility of online content, most users perceived it as intrinsically credible. This could be due to the fact that they viewed the application as a test material for thesis. But aligning with users' disinterest towards verifiability feature

of PSD model, a more realistic assumption is that users simply perceive verification tools as 'credible' systems. This exposes fact-checking systems to another paradox where unreliable fact-checkers could be perceived as trustworthy by users.

Value paradox was more eminent in the system as users struggled to understand the use case of the system instantly. Most of them understood the promise of an interactive fact-checking tool, but some users perceived it as irrelevant and some criticized its time-consuming nature. Again, this could be due to the compliance nature of the application, as news verification is not the primary activity users are seeking. Thus, the application could generate most value for stakeholders whose primary intention is to reduce spread of false news, for example social media companies, news services, governments and non-profit organizations. General users would be most benefitted from the application when it is integrated with their primary activity.

6.1 Limitations

There are several limitations of this study, which should be considered while interpreting the results. One major limitation is regarding the artifact that was developed. The full scope of the artifact was not realized in the final implementation due to resource and time constraints. The architecture of the application and its implementation could have been more rigorous, with individually designed components. Given that the topic was fake news, usability of the system could have been more professional and robust, as the target users were quite broad. Because the event of fact-checking happens in a very short time, it's important that users have a frictionless user experience.

The system itself was not fully functional all the time, as the main content was based on third-party APIs. The technical architecture of the API changed from provider's end during the implementation, which significantly restricted scope of features. Several features such as self-monitoring, praise and rewards were implemented as interactive mockups. However, many participants did not notice during the evaluation which pages were mockups and which were operational. Thus, it could be assumed that users treated mockup pages as part of the main application.

Another limitation of the research is homogenous participants. All participants were students of the same university with similar demographics, which might have affected the results. Most of them were graduate or doctoral students, which means they were fairly educated and had stable income level. Research showed that age, gender, education level, economic status etc. factors are influence people's information sharing behavior (Khan et al. 2019).

Although three experiments were conducted for evaluation, they took place in one single session, in a short timespan. This could lead to bias in collected data. Attitude-forming experiments are usually conducted for longer period, to compare before and after results. This was not possible due to the limited scope of the thesis. Thus, results of the experiments are valid in an overarching way, representing the most dominant features of the application.

7. Conclusion

The main contribution of this research is introducing persuasion as a potential intervention strategy to combat fake news. Fogg's functional triad (1998) and Oinas-Kukkonen et al.'s (2009) persuasive design principles have been applied to a wide range of intervention systems that influence people's behavior to form meaningful habits. Such type of intervention in fact-checking context has been repeatedly suggested in recent scholarly work (Lazer et al., 2017; Lazer et al. 2018; Karduni, 2018; Fernandez and Alani, 2018). Because persuasive systems are inherently designed as behavioral interventions, they could be an ideal strategy to design effective fact-checking systems, just like the way they are used in fitness applications.

PSD analysis in section 2.4.2 showed that current fact-checking tools lack the persuasive features required to be qualified as an 'intervention'. Furthermore, no significant study was found during the literature review which explored the promise of persuasion in fact-checking context. This study tried to bridge this gap by developing Fact-checker, a fact-checking application based on persuasive principles.

The application was developed with general internet users in mind. Theoretical background was adopted from both cognitive psychology and persuasive technology. Scholars from communications and psychology fields suggested cognitive cues to minimize effect of fake news, which were combined with persuasive context. A set of persuasive principles were applied to design it as a behavioral intervention. Goal of the intervention was to encourage users into regular fact-checking, thus forming fact-checking behavior. The system was evaluated with potential users to measure efficiency and persuasiveness, which returned positive results and several interesting insights. One major finding from results was that social persuasive features have strong potential to change people's behavior towards fake news. Another interesting finding was promise of a persuasive gatekeeping mechanism to prevent misinformation sharing.

Future research should focus on identifying cognitive and psychological theories that are most relevant to fact-checking. A number of studies identified the psychological roots of fake news, but much work needs to be done on the mitigation side. Comparing with the research on fake news itself, mitigation techniques are still sparsely studied. More interdisciplinary research is required that would study the mitigation part holistically. Another area for future research could be studying persuasive applications from other fields and adapt their key features in fact-checking context. How fact-checking as a domain is different from others should be also noted while conducting such research, as fake news is more general and mass level than other persuasive fields. Social support persuasive features should be studied further with more detailed system functions.

In summary, fake news evolved a lot in the last decade, and will continue to do so. With the rising popularity of internet and social media, it could be one of the biggest threats on information we have seen. Researchers from interdisciplinary disciplines are developing tools and methodologies to prevent it, but the actual impact of these on mass users is still very limited. It should be noted that victims of fake are general internet users, with or without any technical or analytical skills. Many of these users are not used to with sophisticated systems and graphical representations. A large number of fact-checking tools developed in academia still overlook this fact and builds fact-checking systems tailored for skilled individuals. Thus, solutions targeted for mass audience should priority simplicity and ease of use over computational supremacy.

References

Allcott, H., and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-36.

Apple (n.d.). Human Interface Guidelines. Retrieved from https://developer.apple.com/design/human-interface-guidelines.

Dang, A., Moh'd, A., Milios, E., & Minghim, R. (2016, June). What is in a rumour: Combined visual analysis of rumour flow and user activity. *In Proceedings of the 33rd Computer Graphics International* (pp. 17-20). ACM.

Barnum, C. M., & Palmer, L. A. (2010, April). More than a feeling: understanding the desirability factor in user experience. *In CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4703-4716). ACM.

Benedek, J., & Miner, T. (2002). Measuring Desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of Usability Professionals Association*, 2003(8-12), 57.

Bourgonje, P., Schneider, J. M., & Rehm, G. (2017, September). From clickbait to fake news detection: an approach based on detecting the stance of headlines to articles. *In Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism,* (pp. 84-89).

Brandtzaeg, P. B., and Følstad, A. (2017). Trust and distrust in online fact-checking services. *Communications of the ACM*, *60*(9), 65-71.

Brandtzaeg, P. B., Følstad, A., and Chaparro Domínguez, M. Á. (2018). How Journalists and Social Media Users Perceive Online Fact-Checking and Verification Services. *Journalism Practice*, *12*(9), 1109-1129.

Caldiera, V. R. B. G., & Rombach, H. D. (1994). The goal question metric approach. *Encyclopedia of software engineering*, 528-532.

Camburn, B., Viswanathan, V., Linsey, J., Anderson, D., Jensen, D., Crawford, R., ... & Wood, K. (2017). Design prototyping methods: state of the art in strategies, techniques, and guidelines. *Design Science*, 3.

Casal, J. A., Tubio, O. D., Vazquez, R. G., Fernandez, M. L., & Yanez, S. R. (1998). Formalising the software evaluation process. In Proceedings SCCC'98. *18th International Conference of the Chilean Society of Computer Science (Cat. No. 98EX212)* (pp. 15-24).

Chen, X., & Sin, S. C. J. (2013, November). 'Misinformation? what of it?': motivations and individual differences in misinformation sharing on social media. *In Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries* (p. 104). American Society for Information Science.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Diakopoulos, N., De Choudhury, M., & Naaman, M. (2012, May). Finding and assessing social media information sources in the context of journalism. *In Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2451-2460). ACM.

Endsley, M. R. (2018). Combating information attacks in the age of the internet: new challenges for cognitive engineering. *Human factors*, 60(8), 1081-1094.

F Brown, B., Bødker, S., and Höök, K. (2017). Does HCI scale: scale hacking and the relevance of HCI. *Interactions*, 24(5), 28-33.

Fernandez, M., and Alani, H. (2018). Online Misinformation: Challenges and Future Directions. In *Companion of the The Web Conference 2018 on The Web Conference 2018* (pp. 595-602).

Finn, S., Metaxas, P. T., & Mustafaraj, E. (2014). Investigating rumor propagation with TwitterTrails. *arXiv preprint arXiv:1411.3550*.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1), 298-320.

Flux (2014). Retrieved from https://facebook.github.io/flux

Fogg, B. J. (1998). Persuasive computers: perspectives and research directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 225-232). ACM Press/Addison-Wesley Publishing Co.

Fogg, B. J., & Iizawa, D. (2008). Online persuasion in Facebook and Mixi: A cross-cultural comparison. *In International conference on persuasive technology* (pp. 35-46). Springer, Berlin, Heidelberg.

Forget, A., Chiasson, S., Van Oorschot, P. C., & Biddle, R. (2008). Persuasion for stronger passwords: Motivation and pilot study. *In International Conference on Persuasive Technology* (pp. 140-150). Springer, Berlin, Heidelberg.

Garlan, D., & Shaw, M. (1993). An introduction to software architecture. *Advances in software engineering and knowledge engineering,* 1(3.4).

Garlan, D. (2014). Software architecture: A travelogue. *Proceedings of the on Future of Software Engineering*, 29-39.

Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013, May). Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. *In Proceedings of the 22nd international conference on World Wide Web* (pp. 729-736). ACM.

Hamari, J., Koivisto, J., and Pakkanen, T. (2014, May). Do persuasive technologies persuade? A review of empirical studies. *In International conference on persuasive technology* (pp. 118-136). Springer, Cham

Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., ... & Sable, V. (2017). Claimbuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment*, 10(12), 1945-1948.

Hecht, B., Terveen, L., Starbird, K., Shneiderman, B., and Golbeck, J. (2017, May). The 2016 US election and HCI: Towards a research agenda. *In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1307-1311). ACM.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems*, 28(1), 75-105.

Ibrahim, N., Shiratuddin, M. F., & Wong, K. W. (2013). Persuasion techniques for tourism website design.

Jensen, J. F. (1997). Interactivity: Tracking a new concept in media and communication studies. [Sl]: *Aalborg University. Department of Mechanical Engineering*.

Jin, Z., Cao, J., Zhang, Y., Zhou, J., & Tian, Q. (2017). *Novel visual and statistical image features for microblogs news verification. IEEE transactions on multimedia*, *19(3)*, 598-608.

Jokela, T., Iivari, N., Matero, J., & Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11. *In Proceedings of the Latin American conference on Human-computer interac*tion (pp. 53-60). ACM.

Karduni, A. (2019). Human-Misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information. *arXiv preprint arXiv:1903.07136*.

Khan, M. L., & Idris, I. K. (2019). Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behavior & Information Technology*, 1-19.

Kumar, K. K., & Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences*, 4(1), 14.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., and Schudson, M. (2018). The science of fake news. *Science*, 359(6380), 1094-1096.

Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K.et al. (2017). Combating fake news: An agenda for research and action. *Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy*, 2.

Lehto, T., Oinas-Kukkonen, H., & Drozd, F. (2012). Factors affecting perceived persuasiveness of a behavior change support system.

Lehto, T., & Oinas-Kukkonen, H. (2015). Explaining and predicting perceived effectiveness and use continuance intention of a behaviour change support system for weight loss. *Behaviour & Information Technology*, 34(2), 176-189.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106-131.

Lobodzinski & I. Tomek (Eds.), *Proceedings of WebNet 97*, (pp.185-204) Charlottesville, VA.

Lyons, T. (2017, December 20). Replacing Disputed Flags with Related Articles. *Facebook Newsroom*. Retrieved from https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/

Kaplan, B., & Maxwell, J. A. (2005). Qualitative research methods for evaluating computer information systems. *In Evaluating the organizational impact of healthcare information systems (pp. 30-55)*. Springer, New York, NY.

Khan, M. L., & Idris, I. K. (2019). Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behaviour & Information Technology*, 1-19.

Nguyen, A. T., Kharosekar, A., Krishnan, S., Krishnan, S., Tate, E., Wallace, B. C., & Lease, M. (2018). Believe it or not: Designing a Human-AI Partnership for Mixed-Initiative Fact-Checking. *In The 31st Annual ACM Symposium on User Interface Software and Technology* (pp. 189-199). ACM.

Nuseibeh, B., & Easterbrook, S. (2000, May). *Requirements engineering: a roadmap*. In Proceedings of the Conference on the Future of Software Engineering (pp. 35-46). ACM.

Oeldorf-Hirsch, A., & Sundar, S. S. (2010). Online photo sharing as mediated communication. *In annual conference of the International Communication Association*, Singapore.

Oh, J., & Sundar, S. S. (2015). How does interactivity persuade? An experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *Journal of Communication*, 65(2), 213-236.

Oinas-Kukkonen, H., and Harjumaa, M. (2008). Towards deeper understanding of persuasion in software and information systems. In *Advances in computer-human interaction, 2008 first international conference on* (pp. 200-205). IEEE.

Oinas-Kukkonen, H., and Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems*, *24*(1), 28.

Oinas-Kukkonen, H. (2013). *Humanizing the web: change and social innovation*. Springer.

Orji, R., and Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health informatics journal*, 24(1), 66-91.

Osatuyi, B., and Hughes, J. (2018, January). A Tale of Two Internet News Platforms-Real vs. Fake: An Elaboration Likelihood Model Perspective. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Paul, A., & Nalwaya, A. (2016). *React Native for iOS Development*. Apress.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.

Pierri, F., & Ceri, S. (2019). False News On Social Media: A Data-Driven Survey. *arXiv preprint arXiv:1902.07539*.

Politifact (n.d.), Retrieved from https://www.politifact.com .

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*.

React Native (n.d.), Retrieved from https://facebook.github.io/react-native

Resnick, P., Carton, S., Park, S., Shen, Y., & Zeffer, N. (2014). Rumorlens: A system for analyzing the impact of rumors and corrections in social media. *In Proc. Computational Journalism Conference* (pp. 10121-0701).

Rubin, V., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake news or truth? using satirical cues to detect potentially misleading news. *In Proceedings of the second workshop on computational approaches to deception detection* (pp. 7-17).

Rubin, V. L., Chen, Y., & Conroy, N. J. (2015, November). Deception detection for news: three types of fakes. *In Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 83). American Society for Information Science.

Shandwick, W., & Tate, P. (2016). A Real Plague: Fake News. *Civility in America: A nationwide survey*. Retrieved from https://www.webershandwick.com/uploads/news/files/A_Real_Plague_Fake_News.pdf

Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., & Liu, Y. (2019). Combating Fake News: A Survey on Identification and Mitigation Techniques. *arXiv preprint arXiv:1901.06437*.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592, 96-104*.

Shih, L. H., & Jheng, Y. C. (2017). Selecting persuasive strategies and game design elements for encouraging energy saving behavior. *Sustainability*, 9(7), 1281.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

Soll, Jacob (2018, December 18). The Long and Brutal History of Fake News. *Politico.* Retrieved from https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535

Stenchel, Mark. (2017, February). International Fact-Checking Gains Ground, Duke Census Finds. *Duke Reporters' Lab.* Retrieved from https://reporterslab.org/international-fact-checking-gains-ground/.

Stibe, A., & Larson, K. (2016). Persuasive cities for sustainable wellbeing: quantified communities. *In International Conference on Mobile Web and Information Systems* (pp. 271-282). Springer, Cham.

Sundar, S. S., Knobloch-Westerwick, S., & Hastall, M. R. (2007). News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology*, 58(3), 366-378.

Sundar, S. S., Xu, Q., & Bellur, S. (2010). Designing interactivity in media interfaces: A communications perspective. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2247-2256). ACM.

Tandoc Jr, E. C., Lim, Z. W., and Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153.

The Global Risks Report 2018 (2018). *World Economic Forum*. Retrieved from http://reports.weforum.org/global-risks-2018/digital-wildfires/

Gupta, A., Kumaraguru, P., Castillo, C., & Meier, P. (2014, November). Tweetcred: Real-time credibility assessment of content on twitter. *In International Conference on Social Informatics* (pp. 228-243). Springer, Cham.

Vaishnavi, V., & Kuechler, W. (2004). Design research in information systems.

Venable, J., Pries-Heje, J., & Baskerville, R. (2012, May). A comprehensive framework for evaluation in design science research. *In International Conference on Design Science Research in Information Systems* (pp. 423-438). Springer, Berlin, Heidelberg.

Wu, L., & Liu, H. (2018, February). Tracing fake-news footprints: Characterizing social media messages by how they propagate. *In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 637-645). ACM.

Wunsch, M., Stibe, A., Millonig, A., Seer, S., Dai, C., Schechtner, K., & Chin, R. C. (2015). What makes you bike? Exploring persuasive strategies to encourage low-energy mobility. *In International conference on persuasive technology* (pp. 53-64). Springer, Cham.

Xu, Q., & Sundar, S. S. (2014). Lights, camera, music, interaction! Interactive persuasion in e-commerce. *Communication Research*, 41(2), 282-308.

Zhang, X., & Ghorbani, A. A. (2019). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*.

Appendix A

Please take a minute to answer the questions below.

- I often come across fake news/ hoaxes while browsing internet.
    1. Strongly disagree
    2. Disagree
    3. Neutral
    4. Agree
    5. Strongly Agree
- I often try to check facts of a news when they look suspicious
    1. Strongly disagree
    2. Disagree
    3. Neutral
    4. Agree
    5. Strongly Agree
- I am familiar with at least one of the following fact-checking websites/ tools: Snopes, PolitiFact, FactCheck.org, TruthOrFiction, Hoax-slayer, Facebook Related articles, TweetCred, Hoaxy, ClaimBuster
    1. Strongly disagree
    2. Disagree
    3. Neutral
    4. Agree
    5. Strongly Agree
- I try to stay updated with recent news trends
    1. Strongly disagree
    2. Disagree
    3. Neutral
    4. Agree
    5. Strongly Agree
- How do you verify truthfulness of news that look suspicious?
    1. Check fact-checking websites
    2. Google search
    3. Asking friends/ peers
    4. Own analysis
    5. I don't verify usually
- Please write down topic of a fake news you came across recently (you don't need to remember the exact headline, just write the theme)

    _____

Would you be interested to participate in a short interview regarding a fact-checking tool we are developing?

1. Yes
2. No
3. If yes, please leave your contact: _____

Appendix B

Before using the application:

- Do you often come across fake news?
- Are you familiar with any fact-checking tool?

After using the application:

- What do you think of the application?
- Have you used any similar tool before?
- What do you think of the source-checking feature?
- What do you think of the statistics dashboard?
- What do you think about the virtual badges and rewards?
- You can see how many people have checked sources and rated the news. What do you think of this feature?
- Which features in the app did you like most?
- What do you find confusing in the app?
- Which features would help you to check facts more frequently?
- Would the app encourage you to be more conscious about fact-checking?
- Do you think it would affect your fact-checking behavior?

Appendix C

Complete set of 118 product reaction cards

Accessible	Creative	Fast	Meaningful	Slow
Advanced	Customizable	Flexible	Motivating	Sophisticated
Annoying	Cutting edge	Fragile	Not Secure	Stable
Appealing	Dated	Fresh	Not Valuable	Sterile
Approachable	Desirable	Friendly	Novel	Stimulating
Attractive	Difficult	Frustrating	Old	Straightforward
Boring	Disconnected	Fun	Optimistic	Stressful
Business-like	Disruptive	Gets in the way	Ordinary	Timeconsuming
Busy	Distracting	Hard to Use	Organized	Time-Saving
Calm	Dull	Helpful	Overbearing	Too Technical
Clean	Easy to use	High quality	Overwhelming	Trustworthy
Clear	Effective	Impersonal	Patronizing	Unapproachble
Collaborative	Efficient	Impressive	Personal	Unattractive
Comfortable	Effortless	Incomprhensible	Poor quality	Uncontrollable
Compatible	Empowering	Inconsistent	Powerful	Unconventional
Compelling	Energetic	Ineffective	Predictable	Understandable
Complex	Engaging	Innovative	Professional	Undesirable
Comprehensive	Entertaining	Inspiring	Relevant	Unpredictable
Confident	Enthusiastic	Integrated	Reliable	Unrefined
Confusing	Essential	Intimidating	Responsive	Usable
Connected	Exceptional	Intuitive	Rigid	Useful
Consistent	Exciting	Inviting	Satisfying	Valuable
Controllable	Expected	Irrelevant	Secure	
Convenient	Familiar	Low Maintain	Simplistic	

Appendix D

- Fact-checker helps me to check authenticity of recent online claims

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker helps me to monitor my fact-checking statistics

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker provides me relevant feedback

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker provides incentives/ rewards to check authenticity of news regularly

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker notifies me about potentially fake news

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker is credible

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker shows claims from authentic sources

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker is intuitive and easy to use

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker could influence my fact-checking behavior

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree

- Fact-checker encourage me to think about authenticity before sharing a news

	1	2	3	4	5	
Strongly disagree	[ ]	[ ]	[ ]	[ ]	[ ]	Strongly agree