Instagram Happiness Index

Do people portray themselves happier than they really are? Ranking Countries on Happiness Based on Instagram Hashtag Analysis

Thiago Elid Della Monica
Graduation Thesis, August 2018
Media Technology MSc Program, Leiden University
Supervisors: Peter van der Putten and Ruut Veenhoven
thiago.elid@gmail.com

ABSTRACT

Reflecting on social media behaviour, the primary motivation is to analyse user's Instagram persona by collecting, measuring and interpreting the data about happiness on Instagram. The main idea is to analyse how people portray happiness in the online photo-video-sharing social network based on the Geneva Emotion Wheel. Having the idea of developing a brand-new index in mind, the exploratory work consists of taking the top 11 languages by number of native speakers and make a list of countries that speak those languages. As a result, it is possible to build and rank a full list of countries and then compare the Instagram Happiness Index (IHI) to other accepted metrics for happiness such as the most reliable and prestigious of them, the World Happiness Report (WHR) and by running a correlation test, verify whether there is any association between them.

KEYWORDS: Happiness Index, Geneva Emotion Wheel, Sentiment Analysis, Social Networks, Instagram, Hashtags.

1. INTRODUCTION

In today's connected world, social media is a major channel of communication and an important part of daily life. There are so many positive aspects – not only does social media allow us to easily connect with loved ones and catch up with friends, it also provides entertainment, breaking news and hottest trends from virtually anywhere around the globe. At the same time, there is something that is somewhat negative about the way a particular aspect of the Internet works. People are constantly trying to live a parallel life in social media networks with things happening solely behind the screen and in which sometimes don't truly represent their real life.

This is particularly concerning since Instagram has announced they reached a 700 million monthly active user milestone in April 2017, which grew from 600 million in only four months, meaning it accumulated its latest growth at around the same pace. The huge population of users, make it inevitable to think that within the social network, especially Instagram, people tend to shape their personalities into something that becomes more attractive and exciting than it really is. People seem to create a persona that does not necessarily represent them as they are for real which eventually change the way how they are perceived by others.

A few experiments regarding the persona have already been made with Instagram as a background tool of social behaviour. As we are going to see further ahead, people conduct non-professional tests and sometimes even paid studies aiming to shed light on the topic, yet no study has looked at the platform from its standpoint of happiness and how much of it can be only a real-life overvaluation.

For that reason, this study aims to answer the following research question: Do people portray themselves happier than they really are? And although this is the main subject of the paper, other interesting contrasts will also be seen such as a comparison of the Instagram Happiness Index (IHI) with other accepted metrics for happiness such as the World Happiness Report (WHR) made by the United Nations as well as a series of comparisons of the data by hashtags, likes, comments and last but not least by continent.

We will also look more closely at three factors that have been long studied by science, such as the reward effect, the role of dopamine and selective perception and how these three factors have direct bond connections to the way users make use of social media platforms in these times.

In order to make this paper feasible and reliable from a scientific point of view we also looked at previous studies like the World Happiness Report, the gold standard index in happiness researches and the Geneva Emotional Wheel.

The World Happiness Report was used only as a basis for comparison. The ranking is extremely relevant these days, but it doesn't seem to necessarily make people happier, nor the index is capable of measuring the real happiness of their citizens in a given country.

As for the Instagram Happiness Index, it was freely inspired

by the Geneva Emotional Wheel study, but we took a more technical and data-driven approach in our research compared to other papers that equally use the GEW to develop in depth analysis about emotional reactions and interactions.

With the Instagram Happiness Index, we aim to get a particular view, based on the data collection, on how users portray happiness on Instagram and thus, see a different perspective from the World Happiness Report. Therefore, we believe this paper becomes not only suitable but also relevant for the community since it may indicate what is not being seen yet by researchers and the society itself.

But before heading for an in-depth exploration of the study it is wise to give an overview of the paper and thus better address the audience in order to briefly clarify the concept on a top level. The paper starts looking at studies on the field of happiness, behavioural science and social media connection. Also, we go through studies like Instagram persona, dopamine, reward effect and selective perception.

For the data collection, a breakdown of each of the procedures for crawling user's data from Instagram that resulted in the Instagram Happiness Index, is as follows. Firstly, 20 words were selected from the GEW study. Secondly, a list of the top 11 languages in the world by number of native speakers were chosen resulting in a list of 220 words / hashtags. Thirdly, it was necessary to convert post's geolocation (latitude and longitude) into a country's name. Lastly, it was required to match either hashtags and countries and finally export all the database into a table with 248 countries in total, which is oddly way more than the list of 206 sovereign states of the United Nations.

Wrapping up the paper, the study portrays a new look of happiness on Instagram by analysing the data and providing the reader with a brand-new ranking and a new view on the topic with the Instagram Happiness Index.

2. RESEARCH MOTIVATION

Whenever we think about the purpose of social media networks, especially Instagram, we think of connecting people as its foremost idea. However, as we are going to see, some say this is just the surface of the discussion. Actually, people are constantly creating new public identities aiming to change what they really are to what they desire to be online. Luckily, there are a few experiments that undergo this notion of filtering the reality and what users put out to the world to see it is just to pretend they are all happy people.

2.1 THE HAPPINESS EFFECT

In 2016, Dr. Cal Newport gave a deliberately provocative TEDx Talk titled 'Quit Social Media'. The theme of the event was 'Visions of the Future' and he said that his vision of the future was one in which many fewer people use social media. His claim was even though he has never had a social media account he is happier and more successful professionally. ("Quit Social Media 01:14-02:21").

A similar manner, despite having hundreds of thousands of likes on Instagram Essena O'Neill, a young Australian social media model and millennial influencer, has announced in a 2015 YouTube video that she was quitting social media because it had caused her to constantly compare herself to others and measure her self-worth by the number of likes her posts would get. ("Why I Really Am Quitting").

Similarly, the book "The Happiness Effect: How Social Media Is Driving a Generation to Appear Perfect at Any Cost", by the author Donna Freitas, displays a survey with 800 students of whom she met and interviewed about 200 of them in person, and she compiled it all into a collection of insightful interviews on the social media habits of college students (250-268).

As already mentioned, social media was initially invented and used so that we could better connect with people. But you would be surprised just how disconnected we actually have become because of the manipulation of social media. The book does not only address the issues that come up when talking about Facebook, Twitter or Instagram, it also provides a closer look on how users could possibly make use of social media networks in a much better way without hurting others and most importantly ourselves and our own self-esteem and image. The book tries to lead the way for the readers that perhaps the healthiest thing to do right now after all, is to focus on your real self and not who you would like to be on social media networks i.e. your social media persona.

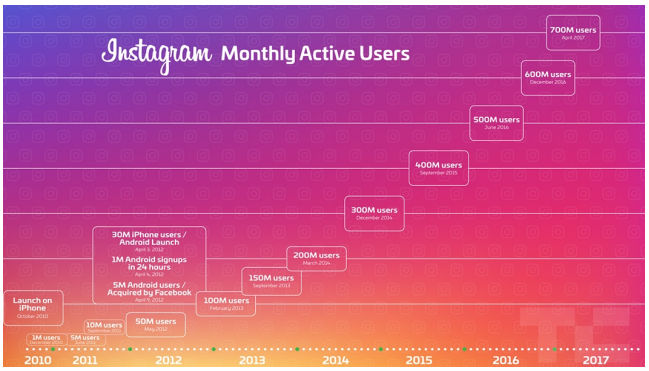

Fig. 1: A breakdown of how long it took Instagram to hit 700 million users. From Constine, Josh. "Instagram's Growth Speeds Up as It Hits 700 Million Users." Techcrunch, 26 Apr. 2017, techcrunch.com/2017/04/26/instagram-700-million-users/

Nonetheless, quitting social media and focusing on your real self doesn't seem to be a trend at least for now. Rather than that, product managers, user experience designers and software engineers behind Instagram are always looking

for ways to make online interaction more fluid and intuitive and why not meaningful too. And it seems that all effort is paying off as Instagram's growth rate is actually speeding up. The app has doubled its user base (April 2017), to 700 million monthly actives in two years, fuelled by 'Stories', web signups and better on-boarding on low-end Android phones (see fig. 1).

2.2 INSTAGRAM PERSONA

You may have already noticed that social media's unique purpose has been to communicate and keep in touch with friends from near and far. However, some of us have been using social media networking to shape our personalities into something that is more attractive and exciting to our online social circle. One of the biggest pet peeves on social media networking sites, like Instagram, is the way people, those of whom we know very well, form their identities or lifestyles into something more fascinating. Nowadays, society cares about what the public has to say about their lives, that's the reason why, every Monday, our Instagram's feed ensures that our friends know that they had a phenomenal and remarkable weekend.

A short film about the contrast between real life and what people share on social media called "What's On Your Mind?" went viral in 2014 by focusing on character development and in showing the contrasts between the main character's real life and how he presents it online. The video uses mesmerising narration and editing to build relationships between elements and show their contrasts, as well as to produce funny effects, and close-up shots to convey the character's emotions and deteriorating situation. ("What's On Your Mind?")

We are all guilty of twisting the truth a little bit on social media, but at which point has it gone too far? This question was also raised in another short film for the Internet made by "Ditch the Label", one of the largest anti-bullying charities in the world, dedicated to promoting equality and providing support to young people who have been negatively affected by bullying and prejudice. ("Ditch The Label") In the short video titled "Are You Living an Insta Lie? Social Media Vs. Reality" Instagram's users pretend and post pictures online of an exciting and happy life that they are far away of living in their real life ("Are You Living").

All previous examples have something in common. Once people become social media's users, especially on Instagram, they seem to switch into something they actually are not in real lives. And a small remark should be added here, there is even an Internet slang for that, the acronym IRL, which is often used to let people know you are talking about something in the real world and not in the Internet world ("IRL").

To put the subject of twisting our lives perhaps a bit beyond the limit, the meme displayed in fig. 2 clearly summarises the idea in a nutshell (Wiktionary contributors). Together with all aforementioned examples of happiness in real life versus in the digital world, a research question came to mind: Do people portray themselves happier than they really are? If so, can we put the outcome into perspective and compare it to the World Happiness Report from the United Nations? And to other different metrics for happiness? This is what we are going to find out in the next pages of this research.

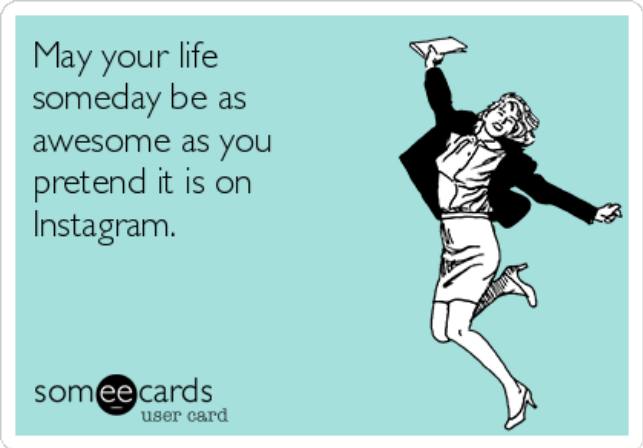

Fig. 2: Meme about perfect life on Instagram. From Kellynnheller, "May Your Life Be Someday Be as Awesome as You Pretend it is on Instagram." Someecards, www.someecards.com/usercards/viewcard/may-your-life-someday-be-as-awesome-as-you-pretend-it-is-on-instagram-58532

2.3 SOCIAL MEDIA EXPERIMENTS

Louise Delage a 25-year-old French social media influencer, who liked spending time with friends, eating at restaurants and being outdoors. Her photos had simple captions like "Chilling with friends" and "Dancing", or sometimes even just an emoji, were hashtagged and received hundreds of likes. She accumulated nearly 65.000 followers in a little over a month. Delage made her last post to Instagram less than 60 days after joining the social media network with a video clip revealing her to be the creation of the advertising agency BETC, and the star of their "Like my addiction" campaign (BETC; "Like My Addiction").

Two to three images were posted each day at periods of high traffic: morning, lunchtime and late at night. Each included up to thirty related hashtags, aiming to reach other Instagram influencers and a bot was created to like and follow carefully chosen accounts on Delage's behalf and thus spread her profile among their own followers thus prompting them to follow her back (Hunt).

In almost every one of her 150 posts, Delage had been pictured with alcohol. The campaign was created for

Addict Aide, an NGO that seeks to raise awareness of alcoholism among young people. The fake Instagram account aimed to show a person that people would want to meet every day but whom we would never suspect of being an addict" (qtd. in Hunt).

By metrics, the campaign was a success, generating media coverage around the world. The final post to Delage's Instagram, the big reveal, has had more than 160.000 views and generated more than 1.000 comments. The same video on YouTube has been viewed on YouTube more than 230.000 times and even nowadays that she is no longer posting any pictures in her profile, she still has more than 110.000 followers.

Another example was Boris Bork. This Moscow tycoon drove fast cars, ate fine food and was frequently seen accompanied by glamorous women. Photos on Instagram portrays a life of luxury with the blessings of an immense wealth and attracted more than 17.000 followers. But all is not quite as it seems, because Boris Bork is not real. Boris Bork is the creation of two friends who set out to prove that you actually do not need too much money to create an online sensation.

Marketing consultant Roman Zaripov, said the idea had come to him and a blogger friend after reading an article about how much it supposedly cost to create a social media star - a figure involving six or seven zeroes.

Convinced they could do it much cheaper, they searched VKontakte, Russia's Facebook, for a grandfather who looked like a fictional millionaire, and found Boris Kudryashov, a pensioner willing to go along with their idea. And they then spent several weekends shooting photographs to turn Mr. Kudryashov, who lives on a modest € 200 a month, into millionaire Boris Bork. "I'm still surprised at how, having spent less than € 1.000 in two months, you can make tens of thousands adults believe in a non-existent person", he said (qtd. in Bell).

Dutch student Zilla van den Born took things a few steps further by faking an entire exotic five-week holiday across South East Asia from the comfort of her own home. From snorkelling in clear blue seas to visiting temples, sunbathing on paradisal beaches to trying new exciting foods, Zilla did it all. Except she didn't. In fact, despite being waved off at the airport by her family, she didn't even leave the Netherlands (Goorwich).

Instead, the 25-year-old graphics student spent the next 42 days at home in Amsterdam with her boyfriend, skilfully photoshopping herself into a variety of envy-inducing holiday snaps, which she then posted on Instagram (Born).

So why did she do it? Apparently, Zilla wanted to demonstrate just how far we can use social media to manipulate the image we put out to the world. She explained: "I did this to show people that we filter and manipulate what we show on social media. We create an ideal world online which reality can no longer meet."

In our last example, Time Magazine asked followers to send images that represented joy to them using the hashtag #TIMEHappiness. They displayed 18 of the most inspiring and imaginative submissions in their website. Images were selected by Time's photo editors; However, it is not surprising that all pictures had more or less always the same background: beaches, lakes, fields, dogs, children and even a roller coaster. Making a long story short: predictable. No need to say there were no photos of someone working in the office as an example ("What Happiness Looks").

After getting to know all the previous examples, it seems clear now, that people are always finding new ways to fake their real lives into something much better and, as a result, something that could possibly make their audience to think "this is the life I want for me." Interestingly, as we are going to see further ahead in the paper, it will become even more clear that being happy online or at least pretend it, turned out to be solely the essential idea behind having an Instagram account today.

3. RESEARCH BACKGROUND

If Instagram is so negative why people are still using it? With a projection to surpass 1 billion users at some point in 2018, why does the number of users continue to grow? Why does the platform have become so addictive? The answer to those questions can be found in prior studies regarding human behaviour, such as the addition to dopamine, the reward effect and their implications on social networks and lastly the selective perception and how it can twist our perception of individual events. These topics might give us some pretty good insights on how filters, overestimation and joy on social media platforms can cause distorted realities on viewers around the world.

3.1 DOPAMINE & THE REWARD EFFECT

Acquiring other companies is a growth strategy used by businesses to bring new products or services, talent, and clients under their company umbrella or to eventually shut down competitors. There is a lot to consider when you are in a position to bring another company into your fold. Yet, Google founder Larry Page boils it all down to one simple question. Page will ask, "Is this something you will use once or twice per day, and does it make your life better?"(qtd. in D'Onfro). We all know Google makes things in a different way, but Page's mindset may tell us a lot about how pivotal a startup company can be to their users and how it can influence their lives.

Google CEO reduces his simple rule to one unique point of view that is to know how many times a user will use the same thing per day, in other words, what is the reward that

the user will have if he uses that same service, product or device daily. And if the reward does exist, then why not to acquire such company. Take for example the Waze app that Google acquired back in 2013 for 1.3 billion dollars. (Cohan).

How many times per day do users make use of the application? And what about the reward? How many hours of traffic jam do users get rid of every single month? Luckily, the reward effect can be explained through science and this is what we are going to see hereafter.

Burrhus Frederic Skinner, commonly known as B. F. Skinner, was one of the most influential of American psychologists. Skinner is regarded as the father of Operant Conditioning by conducting experiments using animals which he placed in a '*Skinner Box*'. He introduced a new term into the law of effect: reinforcement, a behaviour which is reinforced, tends to be repeated or be strengthened while a behaviour which is not reinforced tends to be extinguished or to be weakened (Skinner 59-90).

What he found out, and this becomes relevant and replicable in the world of social media networks and also on Instagram is that rewards at random intervals are more motivating than at fixed intervals. If the rat is conditioned to always receive the food every 5 times that the lever is pressed, and suddenly stops it from receiving it, it quickly stops pressing the lever. If he does it at random intervals, even when the stimulus temporarily vanishes, the rat will be pressing hundreds of times before he gives up. The game gets even better when apart from having random intervals, the reward also becomes random.

This is similar to the effect produced by casino slot machines and the old dopamine game of expectation. We all know there is a reward, and it can be a great reward, but we do not know when we will get it. Each time we try, there is a dopamine discharge and this is a way to generate expectation through reward and this applies to Instagram every time a user make a post and waits to get more and more likes (Jabr).

Professor Judith Donath, director of the Sociable Media Group at the Massachusetts Institute of Technology's Media Lab says, "Social networking provides a series of mini mental rewards that don't require much effort to receive. These rewards serve as jolts of energy that recharge the compulsion engine, much like the frisson a gambler receives as a new card hits the table. Cumulatively, the effect is potent and hard to resist." (qtd in. Rosen 37).

Simon Sinek while speaking about Millennials, in an interview for an episode of Inside Quest with Tom Bilyeu, says, "We are growing up in a Facebook and Instagram world, in other words, we are good at putting filters on things. We are good at showing people that life is amazing even though I am depressed. . . . You have an entire generation growing up with lower self-esteem than previous generations, through no fault of their own, they were dealt a bad hand. Now let's add in technology. We know that engagement with social media and our cell phones releases a chemical called dopamine. That's why when you get a text - it feels good. . . . It's why we count the likes, it's why we go back ten times to see if our Instagram is growing slower, did I do something wrong, do they not like me anymore? The trauma for young kids to be unfriended, because we know when you get, you get a hit a dopamine which feels good. . . . Dopamine is the exact same chemical that makes us feel good when we smoke, when we drink and when we gamble. In other words, it's highly, highly addictive. We have age restrictions on smoking, gambling and alcohol and we have no age restrictions on social media and cellphones." ("Simon Sinek on Millennials in the Workplace 02:30-04:21 ").

In the book "The Power of Habit: Why We Do What We Do in Life and Business", Charles Duhigg narrates how one of the world's largest companies, Procter & Gamble, used habit insights to turn a failing product into one of its biggest sellers. P.&G. is the multinational conglomerate behind a whole range of products, from Downy fabric softener to Bounty paper towels to Duracell batteries and dozens of other household brands. In the mid-1990s, P.&G.'s executives began a secret project to create a new product that could eradicate bad smells. P.&G. spent millions developing a colourless, cheap-to-manufacture liquid that could be sprayed on a smoky shirt, stinky couch, old jacket or stained car interior and make it odourless.

A week passed. Then two. A month. Two months. Sales started small and got smaller. Febreze was a faulty product. The marketing team panicked and decided to conduct several in-depth interviews to figure out what was going wrong. The reason Febreze wasn't selling, the marketers realised, was that people couldn't detect most of the bad smells in their lives. If you live with nine cats, you become desensitised to their smell. If you smoke cigarettes, eventually you don't smell smoke anymore. Even the strongest odours fade out with constant exposure. That's why Febreze was a failure. The product's catchword was hidden from the people who needed it the most. And Febreze's reward, an odourless home, was meaningless to someone who couldn't smell offensive odours in the first place.

The marketers needed to position Febreze as something that came at the end of the cleaning ritual, the reward, rather than as a whole new cleaning routine. The company printed new ads showing open windows and blows of fresh air. More perfume was added to the Febreze formula, so that instead of merely neutralising odours, the spray had its own distinct perfume.

And so Febreze, a product originally conceived as a revolutionary way to destroy odours, became an air freshener used once things are already clean. The Febreze makeover occurred in the summer of 1998. Within two months, sales doubled. A year later, the product brought in $230 million.

Since then Febreze has yielded dozens of spinoffs — air fresheners, candles and laundry detergents — that now account for sales of more than $1 billion a year. Eventually, P.& G. began mentioning to customers that, in addition to smelling sweet, Febreze can actually kill bad odours. Today it's one of the top-selling products in the world (31-55).

Lastly, we will have a closer look in the story of Pepsodent, taken from the same book, which illustrates the same reward principle in a lucid manner. Brushing teeth was important, of course, for maintaining dental hygiene. But in the early 1900's, Pepsodent's inventor couldn't convince people by just saying, "brush everyday". He had to create a new need, even if it was something of his imagination. Then he had an interesting idea, he decided to advertise the Pepsodent toothpaste as a creator of beauty. To do so, he added mint and other refreshing substances so that people could feel that tingling sensation on your mouth and tongue after brushing. That cool tingling sensation is the reward that convinces your mind that using the toothpaste has worked. Companies that make toothpastes deliberately add such chemicals so that you get some sort of signal that the product is working and feel rewarded after a brushing session. In the coming years, the sale of Pepsodent skyrocketed, and brushing teeth using Pepsodent became almost a worldwide habit and their creator made millions in profit (56-59).

3.2 SELECTIVE PERCEPTION

Joseph Carl Robnett Licklider, known simply as J. C. R., was an American psychologist and computer scientist who is considered one of the most important figures in computer science and general computing history. In a study from 1968, he wrote the following on 'The Computer as a Communication Device', "When people do their informational work at the console and through the network, telecommunication will be as natural an extension of individual work as face-to-face communication is now. The impact of that fact, and of the marked facilitation of the communicative process, will be very great, both on the individual and on society. Life will be happier for the on-line individual because the people with whom one interacts most strongly will be selected more by commonality of interests and goals than by accidents of proximity. And communication will be more effective and productive, and therefore more enjoyable" and also "In the end, people will spend much more time in computer-facilitated teleconferences and much less en route to meetings." So whenever we think about being online throughout the day, keeping in touch with friends and making posts multiples times a day it is not only because it makes our lives easier and fun but also because it connects us to the ones we love the most and thus have common interests to share with. (40).

A 2010 study from the Kaiser Family Foundation found a dramatic rise in the amount of time children and teens spend using entertainment media, especially among minority youth. The study looked at the use of TV, computers, video games, music, print, and mobile phones. It found that media use increased by 1 hour and 17 minutes a day over the past five years and it wouldn't be a surprise if these numbers skyrocket nowadays to the point that it becomes a concern in contrast with the past (Rideout).

The report also frames the findings of screen media among kids varies a lot by race and family income level, though not by gender. The report says black children spend about 90 minutes more with screen media each day than white children do and about an hour more than Hispanic children. Vicky Rideout, director of the study said "Despite the educational potential of computers, the reality is that their use for education or meaningful content creation is minuscule compared to their use for pure entertainment. Instead of closing the achievement gap, they're widening the time-wasting gap." (qtd. in Richtel).

A corroborated speech was given by the President of the United States from 2009 to 2016, Barack Obama who told a class of graduating university students in 2010 at Hampton University, Virginia that "Information becomes a distraction, a diversion, a form of entertainment, rather than a tool of empowerment, rather than the means of emancipation" (qtd. in "Barack Obama Criticises").

A football game back in 1951, was the catalyst for extensive research and the introduction of a theory that is now the standard for understanding human perception. Psychologists Albert Hastorf at Dartmouth and Hadley Cantril at Princeton noticed soon after the game the college newspapers of each school began printing stories that seemed to suggest two versions of the truth were in open competition to become the official version of reality.

Hastorf and Cantril took advantage of the phenomena presented to them and recruited a group of Dartmouth and Princeton students and coordinated viewings of the filmed football game and administered questionnaires. Even after watching the game again, test subjects held firm to their beliefs. Dartmouth students still justified their team's behaviour and felt Princeton players were the offenders. Princeton students saw their team as the victims of Dartmouth's unsportsmanlike behaviour. Same game, controlled setting, but totally different experiences.

So even with the best intentions to be impartial, each of us is subject to context and personal experience to our perception of individual events. Then a year later, the two of them published a study that is now considered by many to be the best starting point for talking about cognitive bias and selective perception. (Hastorf).

In a place where everybody seems to be enjoying life all along like Instagram, happiness could turn out to be a very twisted mirror of the real life in the social media network as followers may change their perception of reality because

of the selective perception. If one follows mainly profiles from users that pretend their happy from Monday to Sunday regardless if there are close friends, acquaintances, famous people or social media influencers, there is a great chance of having a terrifically incomplete and unrealistic view of reality throughout time. In addition, especially because people in general are spending more and more time on these sorts of environments solely for entertainment purpose, they tend to get user's posts mixed up with personal life, however they seem to forget that posts are for pure entertainment only and not a mirror (i.e. persona) of what is going on with someone's life behind closed doors.

3.3 HAPPINESS, SCIENCE & SOCIAL CONNECTIONS

If you think it is fame and money what keeps us happy and healthy as we go through life, you're not alone. But according to psychiatrist Robert Waldinger, director of the Harvard Study of Adult Development, you are mistaken. As the director of a 75-year-old study on adult development, Waldinger has unprecedented access to data on true happiness and satisfaction. In his talk, he shares three important lessons learned from a research conducted using data from 268 physically and mentally healthy men since 1938: 1. Social connections are good for us and loneliness kills. 2. The quality of your closest relationships is what matters. 3. Good relationships don't just protect our bodies, they also protect our brains. In short, the main message of the study is: good relationships keep us happier and healthier (qtd. in Shenk; qtd. in Mineo).

So far, academic studies of Instagram's effects on our emotions are quite rare. But it is appealing to extend the application of those effects from Facebook to Instagram studies. Researcher Hanna Krasnova says, "A photo can very powerfully provoke immediate social comparison, and that can trigger feelings of inferiority. You don't envy a news story. You get more explicit and implicit cues of people being happy, rich, and successful from a photo than from a status update" (qtd.in Winter).

Krasnova's research has led her to define what she calls an 'Envy Spiral' peculiar to social media. She says, "If you see beautiful photos of your friend on Instagram, one way to compensate is to self-present with even better photos, and then your friend sees your photos and posts even better photos, and so on. Self-promotion triggers more self-promotion, and the world on social media gets further and further from reality." (qtd.in Winter).

The Royal Society for Public Health (RSPH) is an independent, multi-disciplinary charity dedicated to the improvement of the public's health and wellbeing in the United Kingdom. RSPH and the Young Health Movement have published a new report in early 2017 called *Status of Mind*, examining the positive and negative effects of social media on young people's health ("#StatusOfMind").

The survey was conducted with 1.479 subjects, all 14-24 years old, asking them about five of the most popular social media platforms: Facebook, Instagram, Snapchat, Twitter and Youtube. The aim of the survey was to find out how they feel each of these platforms impacts their health and wellbeing (both positively and negatively) and make comparisons between these platforms. They ranked these issues from -2 (a lot worse), through 0 (no effect) to +2 (a lot better).

The ending result had YouTube in the top of the table as the most positive social network with Instagram coming out in the last place (5th) as the most detrimental to young people's mental health and wellbeing.

4. THE WORLD HAPPINESS REPORT

In order to not only analyse but also evaluate the Instagram Happiness Index we are going to compare the IHI to other gold standard metrics in happiness studies such as the World Happiness Report (WHR), which is one of the most reliable and prestigious of them. Next, there is a brief explanation on how the report is developed every year.

4.1 THE CANTRIL LADDER SCALE METHODOLOGY

Norway is happier. At least that is what the last World Happiness Report, developed by the United Nations in 2017 indicates. The country became happier in the last year. Now, it occupies the 1st place in the ranking while in 2016 the country was the 4th. In the lowest position is Central African Republic, which overthrew Burundi as the saddest or the most 'unhappy' country in the world.

The report gives special attention to the social foundations of happiness for individuals and nations worldwide. The research is done using data that come from the Gallup World Poll and the Cantril Ladder Scale Methodology by crossing economic data from a single question, made to roughly 3000 respondents in each of more than 150 countries investigated, and just one question asking them to evaluate their current lives on a ladder ("Frequently Asked Questions."), as follows:

Imagine a ladder, with steps numbered from zero at the bottom and ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you - with the best possible life being a 10, and the worst possible life being a 0. Which step do you believe you are in? On which step do you think you will stand about five years from now?

Although the top ten countries remain the same as in 2016, there has been some shuffling of places. Most notably, Norway has jumped into first position, followed closely by Denmark, Iceland and Switzerland. These four countries are clustered so tightly that the differences among them are

not statistically significant, even with samples averaging 3.000 underlying the averages. Three-quarters of the differences among countries, and also among regions, are accounted for by differences in six key variables, each of which digs into a different aspect of life.

These six factors are GDP per capita, healthy years of life expectancy, social support (as measured by having someone to count on in times of trouble), trust (as measured by a perceived absence of corruption in government and business), perceived freedom to make life decisions, and generosity (as measured by recent 3 donations). The top ten countries ranked highly on all six of these factors. (Halliwell) For more information, see appendix A.

Ranking of Happiness - World Happiness Report 2017	
1. Norway	146. Yemen
2. Denmark	147. South Sudan
3. Iceland	148. Liberia
4. Switzerland	149. Guinea
5. Finland	150. Togo
6. Netherlands	151. Rwanda
7. Canada	152. Syria
8. New Zealand	153. Tanzania
9. Australia	154. Burundi
10. Sweden	155. Central African Republic

Table 1: Top and bottom list of the World Happiness Report in 2017. From Halliwell, John, et al. World Happiness Report 2017. New York: Sustainable Development Solutions Network, 2017, s3.amazonaws.com/happiness-report/2017/HR17.pdf

Six Key Factors regarding Aspect of Life	
1. GDP per capita	4. Trust
2. Life Expectancy	5. Freedom
3. Social Support	6. Generosity

Table 2: Six Key Factors regarding Aspect of Life. From Halliwell, John, et al. World Happiness Report 2017. New York: Sustainable Development Solutions Network, 2017, s3.amazonaws.com/happiness-report/2017/HR17.pdf

5. DATA ORIENTED APPROACH

With the view to collect data in an assertive way the Geneva Emotion Wheel (GEW) was a fundamental part of this paper acting as a background study in a pre-selected list of words. By doing so it was possible to collect data from 20 different words - both positive and negative - and then turn it into a full list of countries by converting its geolocation into a country's name. All this without having to ask subjects to take part on the survey, as it was done purely by gathering live data from Instagram and converting it into a valid information set to be further analysed, as we are going to see next.

5.1. GENEVA EMOTIONAL WHEEL

The Geneva Emotion Wheel (GEW) is a theoretically derived and empirically tested instrument to measure emotional reactions to objects, events, and situations (Scherer). The GEW has previously been used in a variety of contexts, ranging from managers' affect during decision making to the evaluation of body movements and consumer products research (Bänziger; Beck; Caicedo).

In its original proposed format (see fig. 3), the subject is asked to point out the emotion experienced by choosing intensities for a single emotion or a mix of several emotions out of 20 different emotion categories. They are arranged in a wheel format with the axes being defined by two major dimensions of emotional experience: high control and low control (Y axis) and positive valence and negative valence (X axis). Five degrees of intensity are then proposed, represented by circles of different sizes. In addition, 'None' (no emotion felt) and 'Other' (different emotion felt) options are also provided.

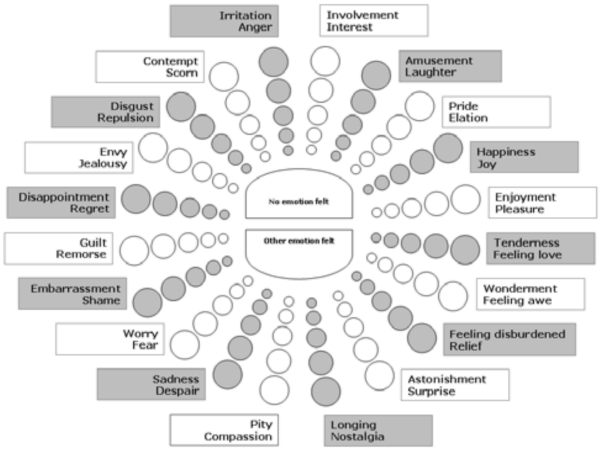

Fig. 3: Original Geneva Emotional Wheel. From Bänziger, Tanja, et al. "The Geneva Emotion Wheel: A tool for the verbal report of emotional reactions." Poster presented at ISRE 149, 2005, pp. 271-294.

The number of emotion families is limited to 10 per quadrant, yielding a total of 40, which seems reasonable considering that the upper limit of the number of basic emotions is often considered to be around 14. The choice of the concrete families was also in large part determined by what are

generally considered to be either basic or fundamental emotions or those frequently studied in the field (Scherer 723).

5.2 LIST OF HASHTAGS

Before starting the data analysis with the words from the Geneva Emotional Wheel, half of the words were ruled out from the original list of 40 words (see fig. 4) in order to make the dataset smaller and to carefully better choose which words could fit into a hashtag analysis on Instagram since it wouldn't be expected to get a lot of data with words such as scorn or elation.

A few words were then chosen to be replaced aiming to get the best possible outcome with only 20 words instead of 40, nonetheless they still fulfil all 4 quadrants from the GEW and at the same time the words were yet pertinent words and all part of the list of 36 semantic categories that index different types of affect-related experiences covering emotions, moods, and other types of transitory affect state from the original GEW study. For more information, see appendix B.

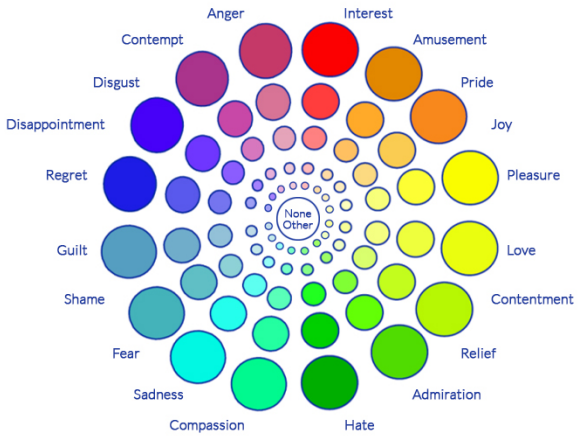

Fig. 4: Modified Geneva Emotional Wheel to fit Instagram's data analysis.

The 20 words selected from the GEW study for the data analysis are the following: 1. interest, 2. amusement, 3. pride, 4. joy, 5. pleasure, 6. contentment, 7. love, 8. admiration, 9. relief, 10. compassion, 11. sadness, 12. guilt, 13. regret, 14. shame, 15. disappointment, 16. fear, 17. disgust, 18. contempt, 19. hate and 20. anger. Above, there is an image from the GEW adjusted to fit the new words previously mentioned. (see fig. 4)

Afterwards, a list of the top 11 languages in the world by number of native speakers were then chosen resulting in a list of 11 arrays, each one of them having all 20 words in their given language ("List of Languages"). The overall number of hashtags was 220. The total number of speakers and their respective languages are seen in fig. 5. For more information, see appendix C.

Number of Native Speakers in Millions (2010)	
1. Chinese = 955	7. Bengali = 205
2. Spanish = 405	8. Russian = 155
3. English = 360	9. Japanese = 125
4. Hindi = 310	10. Punjabi = 100
5. Arabic = 295	11. German = 95
6. Portuguese = 215	TOTAL: 3.220 Billion (47.46%)

Table 3: List of languages by number of native speakers. From Wikipedia contributors. "List of Languages by Number of Native Speakers." Wikipedia, The Free Encyclopedia, 1 May. 2018. en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers.

5.3 SUBJECTS

Participants were not asked to participate in an online survey such as Survey Monkey or Google Forms. Respondents on the other hand were Instagram users. They had some knowledge in one of the 11 most spoken languages in the world. Users also had to have their account turned into a public profile in order to make the posts publicly available, otherwise if the account was private it wouldn't be possible to collect the data.

5.4 COORDINATES AND LOCATIONS

In addition, the geolocation of the post was also necessary aiming to convert latitude and longitude into a country's name, therefore users had to purposely tag a place whenever they were making a post, e.g. Eiffel Tower (see fig. 5), otherwise even though the data was collected the information wouldn't be valid.

Fig. 5: Eiffel Tower tagged in a post by a user and an essential part of the data collected on Instagram. From Guizetti, Fran. Photo of Paris Tour Eiffel. Instagram, 2 Feb. 2017, www.instagram.com/p/BQAjysbjJsU/?taken-by=fran_guizzetti.

5.5 CODING

Amongst all major social media networks, the only one that enables users to properly gather data by collecting hashtags and breakdown all information gathered by country is Twitter. While Facebook does not allow users to get valuable information from hashtags, Instagram on the other hand does not enable users to separate all collected hashtags by country. To overcome such limitation the plan was to work around by writing a code with Instagram's API. The idea behind was to develop a scrapper (i.e. crawler) which is a method

for crawling web sites or social media networks like Instagram and extracting structured data which can be used for a wide range of useful applications, such as data analysis or information processing ("Web Crawler").

In order to collect the data, some steps had to be taken before actually running the code. To make it more readable, a breakdown of the procedure is given below.

After getting all posts saved into the MongoDB database, the next step was to translate both GPS latitude and longitude into the exact location and then turn them into a country's name. Converting coordinates into countries was done using Code Grid Library (ref).

Building the Crawler. (Spider.py)
1. Get all 220 words from the GEW list in 11 languages
2. Replace each word for a #tag
3. Get rid of any stress in Portuguese and Spanish (*diacritics*)
4. Make a request with all posts containing #words
5. Create a JSON file with all posts that were obtained
6. Repeat from step 4 the same request since the last time using *next_url*

Table 4: Breakdown of the code from Instagram's Spider. To check the entire code, see appendix D.

Lastly, after having all the setup ready, it was time to match either tags and countries and finally export all the data into a HTML table.

Turing GPS Location into Countries (Index.js)
1. Extract GPS latitude and longitude from the database
2. Use Code Grid Library (GitHub) to convert it into exact locations
3. Run this script in portions of 1.000 posts each time

Table 5: List of steps on how to get a Country's Name from Code Grid Library / GitHub. For detailed code see appendices E and F.

5.5 RESEARCH LIMITATIONS

Even with the best intentions to be the more accurate as we can in our paper, unfortunately we are aware the study has limitations that will be discussed further ahead in the Future Work section. That being sad, a potential impact could be that at first the study only opens up our eyes with regards to the way Instagram has been used by the users and how distorted the platform becomes from the reality.

6. DATA OUTPUT

Happiness reports are often made out of macroeconomic statistics mixing up numbers about the performance, structure and behaviour of a national economy as a whole. Nonetheless, it leads to a very predictable index, having in general, major developed countries on the top of the list.

The goal then is to create an Instagram Happiness Index that deals with real time data and which is also based on user generated content from Instagram. As a result, we are going to analyse user's Instagram persona by collecting, measuring and interpreting the data about happiness on Instagram.

Exporting the data into a HTML table (Export_Table.js)
1. Get an array of the same word in all 11 languages
2. Count posts with the words from step 1 for each pair of English word and country
3. Get the result with country and tags and all posts gathered
4. Turn all the information gathered into a HTML table

Table 6: Building an array with all tags from the same words. For full code, please refer to appendix G.

The setup run on the 4th of February, 2017 from 04 hours 07 minutes and 37 seconds (GMT-2) until the 5th of February at 19 hours 42 minutes and 22 seconds (GMT-2).

Breakdown of All the Data Collected			
Start:	Feb-04-2017	Hour:	04:07:37 GMT-2
End:	Feb-05-2017	Hour:	19:42:22 GMT-2
Total Duration:		39 hours 34 min 45 sec	
Total Number of Posts:		3.694.850	
Posts without Location:		2.923.922	
Posts with Location:		770.928	
Total Hashtags Collected:		938.541	
Posts with at least 1 Hashtag:		770.928	

Table 7: Breakdown of All the Data Collected.

Users provided data for 39 hours, 34 minutes and 45 seconds, resulting in a sample of 3.694.850 Instagram posts, of which 2.923.922 posts without location and 770.928 posts with location (section 5.4). The total number of hashtags collected in all the posts was 938.541 in which 770.928 posts had at least one of the hashtags mentioned.

A small remark: Instagram doesn't have information such as age or gender about the users. It only offers id, username, profile picture, full name, biography, website, number of posts, number of followers and number of users following.

7. INTERPRETING THE DATA

7.1 BOTS & FAKE ACCOUNTS

In May 2017, a footage has emerged of a giant “click farm” that uses more than 10.000 mobile phones to give likes on social media networks (Jolly). The videoclip shows rows and rows of like making machines all wired to other devices in a factory in China. And there are said to be thousands more phones in the same building all made for the same purpose.
Pretty much anyone can purchase fake followers for just a few dollars and big companies reportedly pay thousands to get their social accounts more likes by using services like this massive “click farm” offers. Experts say digital manipulation is carried out especially in Russia and China but their influence is felt worldwide as we will see in the outcome of the Instagram Happiness Index (IHI).

Unfortunately, Instagram is not the only social media network plagued with fake accounts (Price). In a Brazilian investigation, journalists have found that Twitter, YouTube and Spotify, for instance, are also web platforms where musicians and bands can cheat streaming numbers by buying views, plays, followers and likes. Bots and fake accounts in this case are used to artificially inflate the numbers in a black market that is called “plays market” due to the number of times the play button is pressed (Ortega).

In another article that came out in early 2018, a New York Times investigation found, that the majority of the fake accounts most resemble to real people, with at least 55.000 of them using the names, profile pictures, hometowns and other personal details of real Twitter users. All these accounts belong to customers of an obscure American company named Devumi that has collected millions of dollars in a shadowy global marketplace for social media fraud. Devumi sells Twitter followers and retweets to celebrities, businesses and anyone who wants to appear more popular or display great influence online. Drawing on an estimated stock of at least 3.5 million automated accounts, each sold many times over, the company has provided customers with more or less 200 million Twitter followers, according to the data analysis. (Confessore).

Since the platform is quite aware of such phenomenon, Instagram had decided to crack down hundreds of thousands of them back in December 2014 (Boyer; “#Instapurge”) and they speculated that over 10 million accounts were deleted at the time. (Lorenz).

In 2015, a group of Italian security researchers monitored more than 10.2 million Instagram accounts, including many popular Instagram celebrities such as Taylor Swift and Justin Bieber, and they also purchased 20.000 fake accounts for US$ 100 and then studied their behaviour over the course of a month. Not only did the study find that 7.9% of the accounts in its sample behaved like spam bots, but the researchers also discovered that 29.9% of users appear to be inactive, meaning they post one or fewer photos or videos in a month. (O’Reilly).

7.2 RUSSIA

The chances of having a surprising Instagram Happiness Index with unthinkable countries getting to the top of the list just because their people express themselves online in a very accessible, honest and uncomplicated way seemed to be quite plausible. However, as seen in fig. 6, bots played a very significant role in the outcome of the data collected with Russia getting the 1st place by the number of hashtags gathered. The graph shows that 588.321 out of 938.541 hashtags came from the Russian Federation and it represents 62.68% of all data gathered.

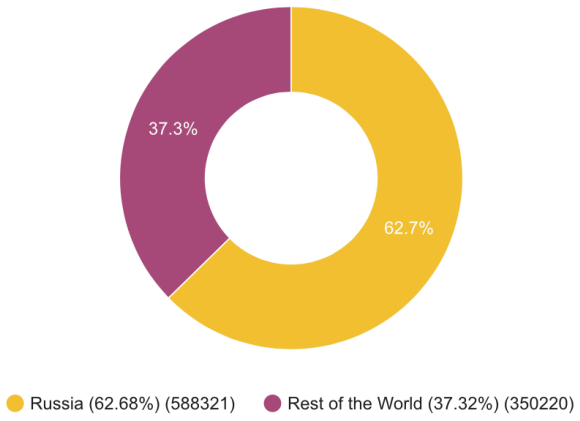

Fig. 6: Number of Hashtags - Russia X Rest of the World (Instagram Happiness Index).

That said, two different interpretations of the dataset were made. One was done only considering data from the Russian Federation whereas the other was done taking into account all remaining 247 countries from the list. Such approach was made so that the data analysed were not inflated by the discrepancy in regards to the volume of the data obtained from Russia.

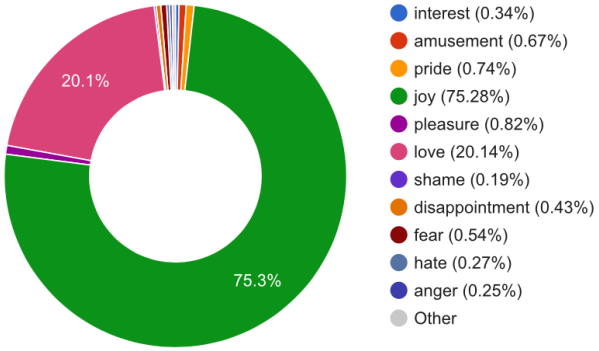

Fig. 7: Hashtags gathered from Russia (Instagram Happiness Index).

Amongst all 20 different hashtags collected two of them got the first places by far. The hashtag joy took the highest place in the podium, while the runner up was the word love. Joy had 442.915 mentions which represents 75.28% of all hashtags collected from Russia while love was used 118.515 times representing 20.14% of all the Russian data. All in all, both words represent 95.42% of the graph above.

When split in two parts, 20 hashtags become 10 words with positive meanings and another 10 with more negative feelings. The positive words are represented by the words interest, amusement, pride, joy, pleasure, contentment, love, admiration, relief and compassion while the negatives are the following sadness, guilt, regret, shame, disappointment, fear, disgust, contempt, hate and anger. Having this in mind, when analysing the Russian data, we get the following result: 98.1% of all hashtags posted were positive words.

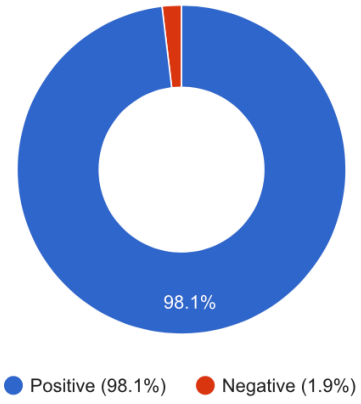

Fig. 8: Russian Hashtags - Positive X Negative (Instagram Happiness Index).

7.3 REST OF THE WORLD

If taken into consideration all remaining 247 countries from the list we have a similar behaviour when it comes to hashtag analysis. Again, amid 20 different words two of them were in the first and second places. Joy was cited 62.69% of the times whereas love got 19.02% of the citations. Altogether, they represent 81.71% of all collected hashtags.

Between positive and negative words, the result was 92.2% positive mentions while negative ones were cited just 7.78% of the times.

8. COMPARING BOTH RANKINGS

As of now, we will take a closer look into the main results obtained from the World Happiness Report (WHR) analysis and then see a breakdown of the ranking aiming to compare

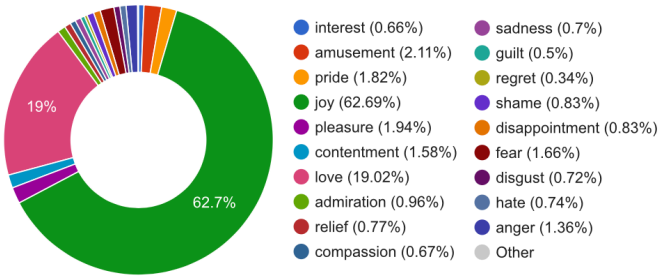

Fig. 9: Hashtags gathered from all remaining 247 countries. (Instagram Happiness Index).

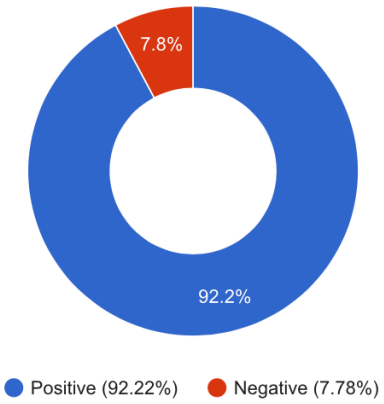

Fig. 10: Rest of the World Hashtags - Positive X Negative (Instagram Happiness Index).

it to the Instagram Happiness Index (IHI) and later, to other accepted metrics for happiness assessment.

8.1 WRH & IHI

The top 10 countries along with the bottom 10 countries from the World Happiness Report were taken in order to compare the results in terms of hashtags of the Instagram Happiness Index.

To get started, the top 10 happiest countries, as seen in table 1 are, 1. Norway, 2. Denmark, 3. Iceland, 4. Switzerland, 5. Finland, 6. Netherlands, 7. Canada, 8. New Zealand, 9. Australia and 10. Sweden and they had, more or less, the same behaviour as seen in fig. 7, with joy (23.87%), love (17.74%) and contentment (7.26%) as the top 3 hashtags (see fig. 11).

If we convert the result into a positive and negative pie chart, the numbers show 73.02% of positive mentions and 26.98% of negative ones (see fig. 12). Although there is a big gap between the total number of hashtags collected in the Instagram Happiness Index, 11.096 hashtags for the top 10 countries versus only 252 hashtags for the bottom 10 countries from the World Happiness Report, a very similar outcome can be seen in the bottom 10 countries from the

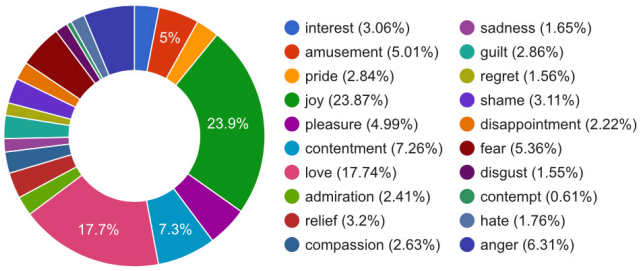

Fig. 11: Top 10 countries from the World Happiness Report and the overall results from the Instagram Happiness Index.

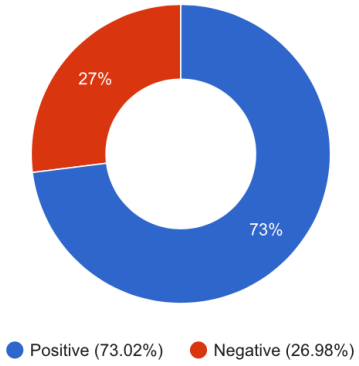

Fig. 12: Top 10 countries from the World Happiness Report and the overall results from the Instagram Happiness Index (Positive X Negative).

World Happiness Report when it is compared to the results of top 10.

Again, the top 3 words in number of hashtags are joy (30.56%), love (15.48%) and contentment (7.54%) as seen in fig. 13. The bottom 10 countries are, 1. Yemen, 2. South Sudan, 3. Liberia, 4. Guinea, 5. Togo, 6. Rwanda, 7. Syria, 8. Tanzania, 9. Burundi and 10. Central African Republic, see table 1.

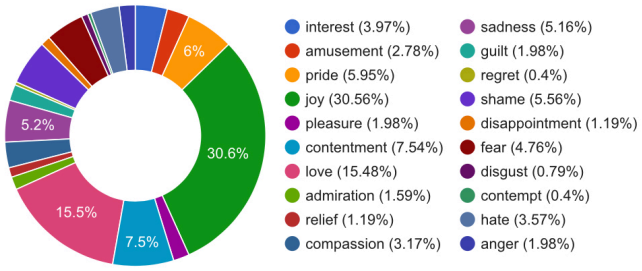

Fig. 13: Bottom 10 countries from the World Happiness Report and the overall results from the Instagram Happiness Index.

When converting the result into a positive and negative pie chart, the numbers show 74.21% of positive mentions and 25.79% of negative ones (see fig. 14). On average, it differs just 1.63% from fig. 12. In terms of hashtag analysis, all 20 countries displayed a very similar behaviour on Instagram.

8.2 IHI & OTHER RANKINGS

Aiming to demonstrate how foreseeable the list of the top 10 countries in the World Happiness Report is, a table together with 5 other rankings is disclosed in table 8. For comparison the table shows the following rankings: Global Workforce Happiness Report (GWHR), Web Index (WI), Human Development Report (HDR), Gross Domestic Product (GDP) and Gross Domestic Product Per Capita (by Purchasing Power Parity) (GDPPP). It is important to stress though that this comparison takes into consideration only their final ranking and not their actual score in each index.

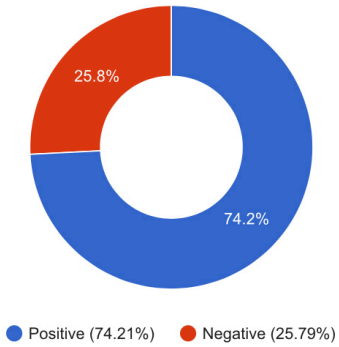

Fig. 14: Bottom 10 countries from the World Happiness Report and the overall results from the Instagram Happiness Index (Positive X Negative).

As seen in table 8, the top 10 countries from the World Happiness Report are placed always in the top of the list of the other 5 rankings, of course taking into consideration its area, population size and so on. Interestingly enough, the second column, IHI which stands for the Instagram Happiness Index, shows the same countries with not so high results. In the next section, we deepen into the countries that turn out to be the top 10 from the Instagram Happiness Index.

9. A BRAND-NEW INDEX

As aforementioned in the section 2. Research Motivation, part of the research question from the Instagram Happiness Index, yet to be answered and measured is: Can we put the outcome into perspective and compare it to the World Happiness Report from the UN with a correlation test?

9.1 MINIMUM SAMPLE SIZE

In order to estimate the minimum sample size for the study (n) the following formula was used to estimate the minimum sample size, in which: n = minimum sample size,

Z = z statistic for a level of confidence (95% level of confidence was used), therefore Z value is 1.96, $\hat{p}$ = expected prevalence of proportion (positive and negative hashtags) and e = sampling error (10%).

For obtaining the expected prevalence of proportion for positive answers ($\hat{p}$) and also obtaining the expected prevalence of proportion for negative answers (1-$\hat{p}$) we considered a conservative value of 0.5 for positive hashtags and 0.5 for negative hashtags meaning that the probability of having either a positive or a negative word in a given post was equal in all cases.

The result for the minimum sample size is n = 96, meaning that only countries with at least 96 hashtags gathered in total will be considered in the study for the final Instagram index.

	WHR	IHI	GWHR	WI	HDR	GDP	GDPPP
Norway	1	189	2	3	1	28	13
Denmark	2	178	1	1	5	34	21
Iceland	3	81		7	9	111	27
Switzerland	4	104	12	18	3	19	15
Finland	5	75	7	2	23	43	33
Netherlands	6	105	6	9	7	17	22
Canada	7	162	23	16	10	10	32
New Zealand	8	144		12	12	53	45
Australia	9	156	24	13	2	13	29
Sweden	10	183	4	5	13	22	25

Table 8: Table comparing the World Happiness Report with 5 other major rankings. Considering only their ranking and not their actual score in each index.

$$n_0 = \frac{Z^2 p(1-p)}{e^2}$$

n	p	(1-p)	α	α/2	Z	e
96	50%	50%	10%	5%	1,96	10%

Fig. 15: Formula and results determining the minimum sample size for the Instagram Happiness Index.

Therefore, in the final exploratory work of the Instagram Happiness Index, we only included a list of 97 countries while the remaining list of 151 countries will be ruled out due to the minimum number of 96 hashtags per country (n = 96). Russia, as aforementioned in section 7.1 and 7.2, will not be considered due to the bots and fake accounts that considerably inflated the number of data collected.

9.2 INSTAGRAM HAPPINESS INDEX

After collecting 938.541 hashtags a brand-new ranking was then built with 248 countries in total. However, as explained in the previous section 9.1 Minimum Sample Size, only 97 countries will be taken into consideration in order to create the new index. The ultimate version of the Instagram Happiness Index can be found in appendices H and I, while the final exploratory version of the Instagram Happiness Index can be found in appendices J and K.

Table 9 shows a short version of the ranking with only the top 10 countries. Georgia, the winning country had 2.917 hashtags gathered overall, being 99.18% of them positive words (blue dots) while just 24 of them being negative words (red dots), it means in numbers that 2.893 were positive hashtags. The top 10 countries are, 1. Georgia, 2. Montenegro, 3. Uzbekistan, 4. Belarus, 5. Kazakhstan, 6. Moldova, 7. Ukraine, 8. Estonia 9. Maldives and 10. Kyrgyzstan. A breakdown of all hashtags from the top 10 countries along with all positive and negative percentages can be seen in table 9.

Top 10 Coutries (IHI)	1. Georgia	2. Montenegro	3. Uzbekistan	4. Belarus	5. Kazakhstan	6. Moldova	7. Ukraine	8. Estonia	9. Maldives	10. Kyrgyzstan	Total (# by word)
interest •	4	1	2	46	163	8	252	5	4	7	492
amusement •	19	1	4	135	834	7	326	9	1	6	1342
pride •	12	2	17	140	286	15	313	7	1	40	833
joy •	1988	605	1485	16495	19121	1149	42856	859	335	1361	86254
pleasure •	22	9	18	189	170	16	675	17	16	18	1150
contentment •	34	0	1	2	3	0	8	0	2	1	51
love •	813	157	593	3957	4325	212	10833	99	175	223	21387
admiration •	0	2	2	1	0	2	26	4	2	0	39
relief •	1	0	1	15	39	1	202	1	1	4	265
compassion •	0	1	0	1	2	0	3	0	0	1	8
sadness •	0	0	0	17	9	2	46	2	2	0	78
guilt •	0	0	0	1	2	0	5	1	0	0	9
regret •	1	0	1	8	11	0	12	2	3	0	38
shame •	0	2	0	38	41	6	145	2	1	1	236
disappointment •	3	3	8	108	101	5	214	1	0	22	465
fear •	9	2	11	73	130	7	174	5	5	5	421
disgust •	0	0	0	3	42	1	78	0	0	0	124
contempt •	0	0	0	7	5	1	32	1	0	0	46
hate •	4	1	3	48	40	2	163	5	0	6	272
anger •	7	1	2	32	59	1	203	2	2	7	316
Total (# by country)	2917	787	2148	21316	25383	1435	56566	1022	550	1702	113826
Positive	2893	778	2123	20981	24943	1410	55494	1001	537	1661	111821
Positive (%)	99.18	98.86	98.84	98.43	98.27	98.26	98.10	97.95	97.64	97.59	98.31
Negative	24	9	25	335	440	25	1072	21	13	41	2005
Negative (%)	0.82	1.14	1.16	1.57	1.73	1.74	1.90	2.05	2.36	2.41	1.69

Table 9: Top 10 countries from the Instagram Happiness Index.

When analysing the top 10 countries from the Instagram Happiness Index, a very similar outcome regarding the number of mentions per hashtag can be seen. Again, the top 2 words in number of hashtags are joy (75.78%) and love (18.79%) as seen in fig. 16.

Once more, when converting the result into a positive and negative pie chart, the numbers show 98.24% of positive mentions and 1.76% of negative words (see fig. 17).

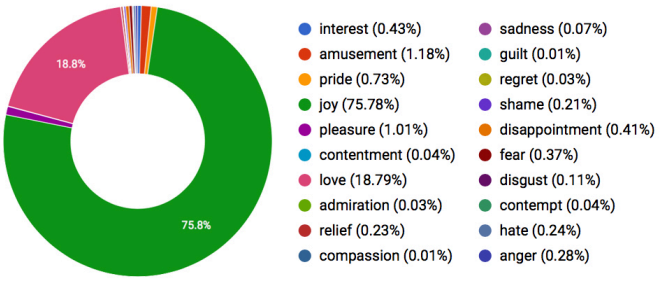

Fig. 16: Top 10 countries from the Instagram Happiness Index (Percentage by Hashtag).

After leaving out countries with less than 96 hashtags, likewise, it was feasible to obtain the bottom 10 countries from the Instagram Happiness Index. Surprisingly, Japan happened to embitter the very last position in the Instagram Happiness Index.

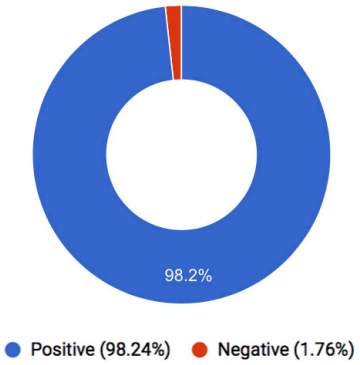

Fig. 17: Top 10 countries from the Instagram Happiness Index (Positive X Negative).

Furthermore, Denmark, Sweden and Norway, three of the top countries from the World Happiness Repost were ranked in the bottom 10 countries from the Instagram Happiness Index. The bottom 10 countries are the following 88. Singapore, 89. The United States of America, 90. Denmark, 91. Pakistan, 92. Sweden, 93. Bosnia and Herzegovina, 94. Norway, 95. Chile, 96. Ireland, and lastly 97. Japan. A breakdown of all hashtags from the bottom 10 countries along with all positive and negative percentages can be seen in table 18.

This time, the bottom 10 countries from the Instagram Happiness Index, shows an interesting outcome with 19 hashtags, each one of them having more than 1%. On the other hand, only 1 hashtag had less than 1%, contempt with 0.6%, as seen in figure 18. The top 3 positive words are joy (15.79%), love (14.47%) and contentment (7.66%). On the contrary, the most negative words used were anger (10.92%), fear (6.76%) and sadness (3.93%). When we convert the result into a positive and negative pie chart, the

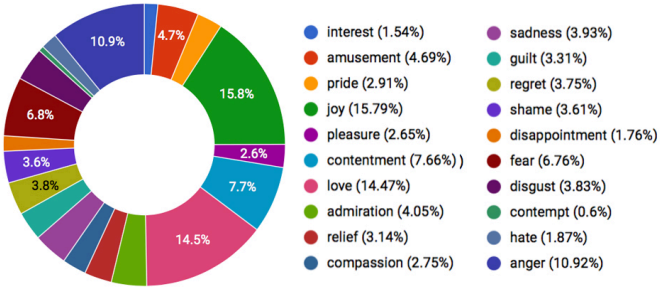

Fig. 18: Bottom 10 countries from the Instagram Happiness Index (Percentage by Hashtag).

Bottom 10 Coutries (IHI)	88. Singapore	89. USA	90. Denmark	91. Pakistan	92. Sweden	93. Bosnia & Herzegovina	94. Norway	95. Chile	96. Ireland	97. Japan	Total (# by word)
interest	27	465	36	4	9	2	29	15	4	130	721
amusement	25	1108	58	7	26	0	5	130	6	826	2191
pride	12	731	4	6	21	2	10	111	13	449	1359
joy	245	4575	107	8	250	41	194	263	53	1638	7374
pleasure	122	681	22	6	25	7	13	149	8	203	1236
contentment	38	872	5	4	300	9	7	65	13	2265	3578
love	112	4488	76	21	117	28	86	348	37	1443	6756
admiration	8	842	5	2	6	2	4	56	11	956	1892
relief	9	799	25	4	9	0	8	74	11	529	1468
compassion	9	1140	9	2	18	3	23	37	7	38	1286
sadness	13	579	10	8	7	3	6	115	6	1087	1834
guilt	12	708	14	1	10	4	10	63	13	710	1545
regret	21	394	3	1	6	0	6	2	5	1313	1751
shame	6	727	6	2	3	1	127	110	8	697	1687
disappointment	40	571	4	2	16	20	6	71	9	82	821
fear	52	1289	82	5	23	23	42	218	16	1406	3156
disgust	17	1434	26	1	9	1	8	247	7	37	1787
contempt	6	243	2	2	3	1	4	1	1	18	281
hate	153	554	5	4	14	2	4	97	7	31	871
anger	7	2185	42	12	375	2	23	60	65	2328	5099
Total (# by country)	934	24385	541	102	1247	151	615	2232	300	16186	46693
Positive	607	15701	347	64	781	94	379	1248	163	8477	27861
Positive (%)	64.99	64.39	64.14	62.75	62.63	62.25	61.63	55.91	54.33	52.37	60.54
Negative	327	8684	194	38	466	57	236	984	137	7709	18832
Negative (%)	35.01	35.61	35.86	37.25	37.37	37.75	38.37	44.09	45.67	47.63	39.46

Table 10: Bottom 10 countries from the Instagram Happiness Index.

numbers unveil the reason why those countries end up in the last places of the index. In short, the bottom 10 countries from the Instagram Happiness Index had just 59.67% of positive hashtags and an impressive 40.33% of negative citations (see fig. 19).

On average, the top 10 countries and the bottom 10 countries differ 38,5% from each other.

10. FUTHER DATA ANALYSIS

10.1 HASHTAGS, LIKES AND COMMENTS

When we take a closer look to the number of likes (heart symbol ♡ on Instagram which also means like) and

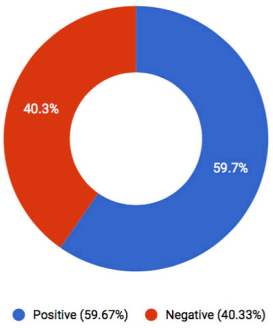

Fig. 19: Bottom 10 countries from the Instagram Happiness Index (Positive X Negative).

comments that each hashtag had in the dataset (table 11), we clearly see a similar behaviour to the numbers found in figure 20. Once again, the first hashtag was joy with 48.859.977 likes (61.05%), the second was love with 16.955.617 (21.19%) and the third was contentment with 3.065.925 (3.83%). The same happened with the number of comments with joy being the first one with 1.338.871 comments (56.62%), the second was love with 510.039 comments (21.57%) and the third one was contentment with 92.288 comments (3.9%).

# Words	Hashtags	Likes	Comments
interest	6185 (0.66%)	445474 (0.56%)	14105 (0.6%)
amusement	19817 (2.11%)	1228953 (1.54%)	38076 (1.61%)
pride	17084 (1.82%)	1467232 (1.83%)	52181 (2.21%)
joy	588398 (62.69%)	48859977 (61.05%)	1338871 (56.62%)
pleasure	18238 (1.94%)	1284863 (1.61%)	42150 (1.78%)
contentment	14854 (1.58%)	3065925 (3.83%)	92288 (3.9%)
love	178468 (19.02%)	16955617 (21.19%)	510039 (21.57%)
admiration	9023 (0.96%)	791538 (0.99%)	25526 (1.08%)
relief	7198 (0.77%)	510113 (0.64%)	18525 (0.78%)
compassion	6278 (0.67%)	471129 (0.59%)	19332 (0.82%)
sadness	6528 (0.7%)	456951 (0.57%)	24539 (1.04%)
guilt	4662 (0.5%)	282615 (0.35%)	14516 (0.61%)
regret	3152 (0.34%)	197072 (0.25%)	10786 (0.46%)
shame	7798 (0.83%)	512675 (0.64%)	26561 (1.12%)
disappointment	7751 (0.83%)	528557 (0.66%)	22408 (0.95%)
fear	15544 (1.66%)	1054603 (1.32%)	43679 (1.85%)
disgust	6746 (0.72%)	431666 (0.54%)	17137 (0.72%)
contempt	1085 (0.12%)	58667 (0.07%)	2159 (0.09%)
hate	6935 (0.74%)	640519 (0.8%)	19549 (0.83%)
anger	12797 (1.36%)	785527 (0.98%)	32042 (1.36%)

Table 11: Table comparing the number of hashtags, likes and comments per word.

The next graph (see fig. 20) shows the peak from the number of likes and the number of comments as aforementioned. As discussed in the entire paper, it is noteworthy that none of the words stood out among the negative results, making all words to obtain more or less the same performance with fear being the word with a slightly better performance in number of likes. Fear had

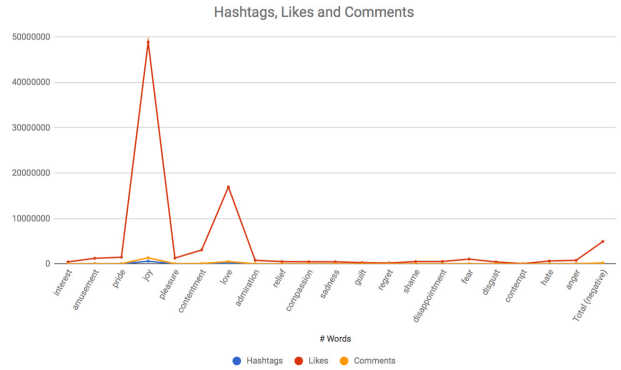

Fig. 20: Line Chart comparing the number of hashtags, likes and comments per word.

1.054.603 likes and 43.679 comments, however it only represents 1.32% of likes and 1.85% of comments. Altogether, negatives words had a sum of 6.18% likes and 9.02% of comments, which corroborate the numbers that were previously analysed in the section 5 of this paper and also shed light on the meme previously seen in fig. 1.

10.2 A COMPARISON BY CONTINENT

Another way to see how the numbers from the Instagram Happiness Index behave is to analyse them by their outcome in all five continents in the world.

Therefore, an extra analysis from the same dataset collected on the 4th of February, 2017 was made in order to obtain a comparison of the numbers per continent.

But before getting into the numbers we got, it is relevant to have a closer look on how the population in the world is distributed by continent (see fig. 21). The graph shows the distribution of global population as of 2017, and about 59.63% of the global population was living in Asia whereas all the rest of the global population was living in all 4 other continents.

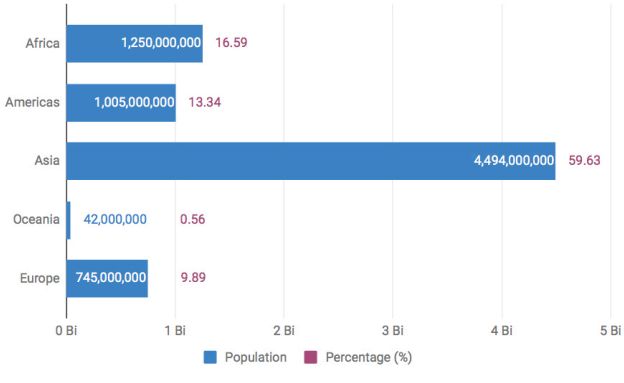

Fig. 21: Graph showing the distribution of the global population in 2017, by continent.

The total world population is around 7.5 billion people on the planet (Population Reference Bureau; DSW). Apart from Asia that has almost 60% of the global population, it is also worth mentioning Europe and Africa with 9.89% and 16.59% of the world population. As table 12 and fig. 22 show, the only hashtag that has a peak in number of mentions is the word joy in Europe which was cited 84.934 times. However, when we compare the number of global population with the data gathered in the dataset we see first, that Asia has slightly less than half the numbers of hashtags (28.17%) if compared to the Asian population in the world (59.63%). Second, Europe has more than 4 times the number of hashtags (42.14%) in comparison to the number of their population in the world (9.89%). Third, Africa has only 1.55% of the whole number of hashtags collected while in terms of global population the continent corresponds to 16.59%.

Distribution of hashtags per continent (*)	Africa	Americas	Asia	Oceania	Europe	Total #
interest	73	1154	1301	61	1577	4166
amusement	116	7862	3616	102	4178	15874
pride	101	6345	2388	87	3800	12721
joy	2379	13461	44404	305	84934	145483
pleasure	171	5472	3309	135	4324	13411
contentment	359	3633	7361	280	3136	14769
love	879	15885	15494	635	27060	59953
admiration	156	5737	1405	118	1436	8852
relief	113	3814	1461	125	1395	6908
compassion	232	4258	884	103	789	6266
sadness	176	2427	2346	52	1354	6355
guilt	68	2106	1281	112	1090	4657
regret	39	644	1840	100	444	3067
shame	68	3936	1011	80	1574	6669
disappointment	41	2331	1494	90	1243	5199
fear	242	5881	3536	136	2547	12342
disgust	42	3427	495	45	2029	6038
contempt	26	436	107	25	239	833
hate	58	1929	1172	55	2145	5359
anger	98	5044	3757	96	2303	11298
Total #	5437	95782	98662	2742	147597	350220
Total (%)	1.55	27.35	28.17	0.78	42.14	100.00
Global Population (in millions)	1250	1005	4494	42	745	7536
GP (%)	16.59	13.34	59.63	0.56	9.89	100.00
(*) 350220 represents the total number of hashtags collected, excluding Russia						

Table 12: Table comparing the distribution of hashtags, per continent, excluding data from Russia.

With that said, we can assume that such issue may be related to how accessible the internet and the Instagram app is in places like Africa, for instance. Or due to the fact that Instagram has been blocked in mainland China since September, 2014 ("Websites Blocked"). However, most of the features and functions of Instagram has been incorporated across other social media platforms like WeChat and Weibo which are operated by the Chinese duopoly Tencent and Sina Corp respectively. To corroborate that impression, Oceania represents only 0.78% of all the dataset but their population has likewise only 0.56% of the population in the world. Nonetheless, a further study in the field would be necessary in order to correctly state that internet access in Africa or internet censorship and surveillance in China are intrinsically correlated to the number of hashtags collected in this study.

11. CONCLUSION

11.1 RESEARCH QUESTION

Now it is time to analyse the proportion of happiness in the 97 countries of the Instagram Happiness Index, considering 95% of confidence intervals as seen in section 9.1. (see fig. 23)

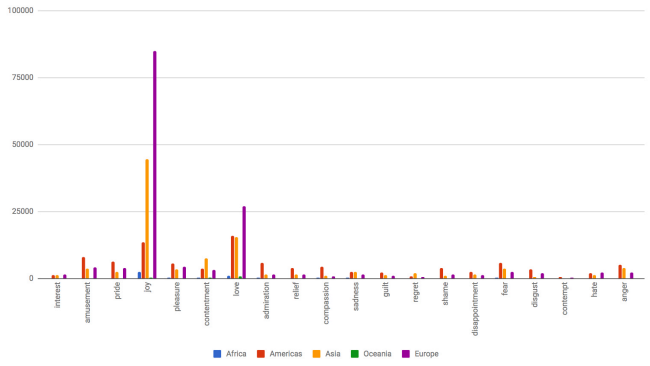

Fig. 22: Column Chart comparing the distribution of hashtags, per continent, excluding data from Russia.

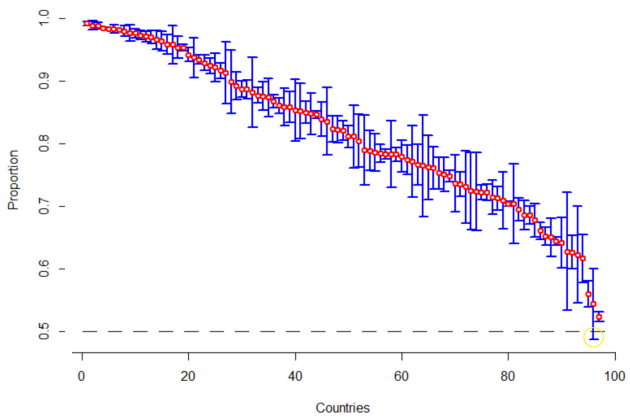

Fig. 23: Confidence interval of 95% for the probability of using positive words according to the ranking of countries.

Only for the penultimate country, Ireland, the lower limit of the confidence interval was < 0.50, indicating that only in this case there is no sample evidence for a significant

difference between the probabilities of using positive and negative words (yellow circle in the lower right corner) (see fig. 23). For all the other countries, the probability of using positive words is significantly higher, since the lower limits of the confidence intervals were always > 0.5. The list with all countries and confidence intervals can be seen in appendix K.

After setting up the discussion and defining the research question: Do people portray themselves happier than they really are?, we can finally answer the research question by saying that, at least for the numbers obtained in the proportion of the happiness analysis (see fig. 23) and according to the comparison seen in fig. 24, there is, as expected (see fig. 2), a greater chance that Instagram users are using the social network in an overestimated way having in most of the time positive hashtags, and consequently, positive posts instead of negative ones. Over 90% of the posts were positive, which shows that they tend to share more positive than negative moments of their lives.

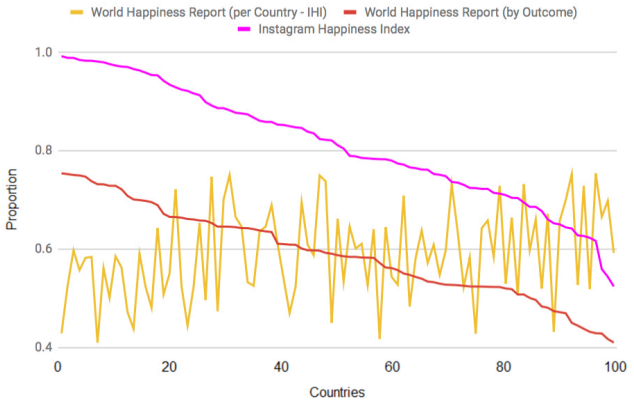

Fig. 24: Comparison between the Instagram Happiness Index X the World Happiness Report considering 95% of confidence intervals.

Additionally, in order to compare the two rankings (the World Happiness Report and the Instagram Happiness Index) a Scatter Plot graph was built (see fig. 25). The scales of the two rankings are represented on each axis. The points indicated in red represent the top 20 countries on the Instagram Happiness Index whereas the blue points are the top 20 of the World Happiness Report.

The points indicated in red have a maximum rating in the Instagram Happiness Index, but at most they reached a 6,5 grade in the World Happiness Report, while the blue points obtained a maximum score in the World Happiness Report and are evenly distributed amongst the ranking range of the Instagram Happiness Index. The only 3 countries that had a good result and therefore performed well in both rankings were Israel, Malta and Thailand (yellow circles in the top right).

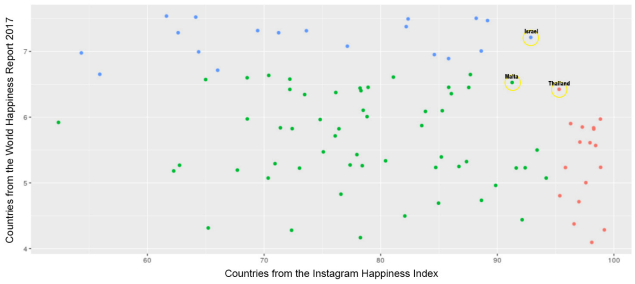

Fig. 25: Dispersion of the top 20 countries in the two rankings studied (axis y = WHR and axis x = IHI).

Additionally, a few other countries had pretty much the same outcome when comparing the results from both rankings: New Zealand, Singapore and Australia. Likewise Canada, Brazil and Argentina had a similar performance even though with a slight different result when looking to the numbers they have got. Apart from that analysis no other pattern could be found in the the dispersion graph (Scatter Plot) when comparing the Instagram Happiness Index and the World Happiness Report. For more information, see appendix L.

11.2 CORRELATION TEST

Lastly, aiming to analyse whether the is a correlation between the two rankings or not, Spearman's rank correlation coefficient was applied ("Spearman's Rank"). Spearman is a nonparametric measure of rank correlation (statistical dependence between the rankings of two variables), whose correlation was estimated by rho = -0.30, indicating a low correlation between the rankings. For the null correlation test (H_0: p value = 0), the p value obtained was equal to 0.9985, or in other words, at the 5% level of significance, we cannot conclude that the association is significant and, therefore, the two rankings are not associated. Then we reject the null hypothesis H_0 and accept the alternative hypothesis H_1 = true.

Spearman's rank correlation rho
data: data.compareihi93$Instagram and data.compare93$whr
S = 169540, p-value = 0.9985
alternative hypothesis: true
rho is greater than 0
sample estimates: rho -0.3064806

Table 13: Output for the Spearman's Rank Correlation in R statistics environment.

It should be stressed that the comparison between the two rankings deserved attention considering the differences between them. The result seen in fig. 24 and table 13, only took into account countries that were both in the Instagram Happiness Index and in the World Happiness Report,

resulting in 93 countries in total. The following countries were ruled out from the Instagram Happiness Index aiming to match the World Happiness Report: Maldives, Monaco, Cuba and Oman. For more information, see appendix M.

So in summary, the Instagram Happiness Index suggests that economic, cultural and technical factors may have influenced the outcome of the research, such as the lack of accessibility to social networks like Instagram in African countries or in countries with authoritarian political regimes. In addition, not only do people seem to portray themselves happier than others on instagram, but also the lack of a relationship with a generally accepted measure such as the World Happiness Report ranking casts further doubts on whether the level of happiness expressed on Instagram is real. . In the next section we will compare the Instagram Happiness Index to other gold standard measure, the Gallup World Poll.

11.3 GALLUP WORLD POLL

The Gallup World Poll (GWP) has been conducted annually since 2006 around the world in more than 160 countries that include 99% of the world's adult population. With some exceptions, all surveys, either telephone or face-to-face, are probability based and nationally representative of the resident population aged 15 and older. The survey includes more than 100 global questions as well as region-specific items. Gallup asks residents from Australia to Pakistan the same questions, every time, in the same way. This makes it possible to trend data from year to year and make direct country comparisons.

Luckily, the Gallup World Poll also contains an independent happiness measure that better aligns with the emotional experience that is communicated on the Instagram Happiness Index, namely a number of questions especially in regards to health and well-being that reflect on how a subject felt yesterday. This comparison with yesterday's affect balance results in a higher quality comparison, not only from a purpose but also from a quantification point of view. The difference is between asking an average citizen of a country how do they feel by anonymous interviewers, versus what Instagram users tell to their network on a daily basis.

Based on the answers of these questions an affect balance score was then calculated by subtracting the percentage of reported negative feelings from the percentage of positive feelings. Following collecting the data from the Gallup World Poll, a new table containing data from all three rankings - the Instagram Happiness Index (2017), the World Happiness Report (2017) and the Gallup World Poll (2016) - was then built (see table 14). For more information, see appendix N. (Veenhoven)

After analysing the data from the top 20 countries from the Gallup World Poll it is possible to see that five countries are too on the top 20 countries of the Instagram Happiness Index, they are: Uzbekistan, Belarus, Kazakhstan, Kyrgyzstan and Thailand. To corroborate four other countries are ranked on the bottom 20 of either the Gallup World Poll and the Instagram Happiness Index: Pakistan, Bosnia and Herzegovina, India and Croatia, as seen in figure 26. On the other hand a remark should be made due to the fact that none of the top 20 countries of the Instagram Happiness Index are listed in the top 20 of the World Happiness Report. The same applies for the bottom 20 in both rankings. For more information, see appendix O.

IHI Rank	WHR Rank	GWP Rank	Countries (Alphabetical Order)	Instagram Happiness Index	World Happiness Resport	Gallup World Poll
3	87	1	Uzbekistan	9.88	5.97	8.40
80	24	2	Singapore	6.50	6.57	8.20
67	31	3	Taiwan, Republic of China	7.22	6.42	8.10
9	74	4	Kyrgyzstan	9.76	5.00	8.05
82	2	5	Denmark	6.41	7.52	7.95
85	1	6	Norway	6.16	7.54	7.85
41	36	7	Bahrain	8.38	6.09	7.85
68	23	8	Mexico	7.22	6.58	7.80
28	3	9	Iceland	8.82	7.50	7.75
43	4	10	Switzerland	8.24	7.49	7.75
25	5	11	Finland	8.92	7.47	7.75
17	30	12	Thailand	9.53	6.42	7.75
4	52	13	Belarus	9.84	5.57	7.75
44	6	14	Netherlands	8.22	7.38	7.70
62	8	15	New Zealand	7.36	7.31	7.70
30	28	16	Panama	8.75	6.45	7.70
5	48	17	Kazakhstan	9.83	5.82	7.70
27	12	18	Austria	8.86	7.01	7.65
87	14	19	Ireland	5.43	6.98	7.65
55	60	20	China	7.74	5.27	7.65
47	57	69	Jordan	8.04	5.34	6.65
79	83	70	India	6.52	4.32	6.65
6	44	71	Moldova	9.83	5.84	6.55
32	63	72	Venezuela	8.67	5.25	6.55
64	69	73	Lebanon	7.30	5.23	6.55
37	56	74	Serbia	8.52	5.40	6.50
2	64	75	Montenegro	9.89	5.24	6.50
84	71	76	Bosnia and Herzegovina	6.23	5.18	6.50
1	84	77	Georgia	9.92	4.29	6.50
11	50	78	Cyprus	9.71	5.62	6.45
71	59	79	Croatia	7.09	5.29	6.30
83	61	80	Pakistan	6.27	5.27	6.30
22	68	81	Greece	9.16	5.23	6.25
42	42	82	Algeria	8.35	5.87	6.20
26	78	83	Egypt	8.86	4.74	6.20
24	75	84	Nepal	8.99	4.96	6.00
16	77	85	Tunisia	9.54	4.81	6.00
19	53	86	Turkey	9.34	5.50	5.60
38	80	87	Iran	8.50	4.69	5.55
45	81	88	Iraq	8.21	4.50	4.80
(*) Countries with data unavailable or invalid due to the minimum sample size in at least one of the three rankings were excluded. (see section 9.1)						

Table 14: Table comparing the top 20 and bottom 20 countries in all three rankings (Instagram Happiness Index X World Happiness Report X Gallup World Poll).

If we analyse the data of both rankings but this time focusing of the grade itself rather than the ranking then we can see that some of the countries had pretty much the same result: Argentina, Australia, Brazil, Canada, Costa Rica, Netherlands, New Zealand, Switzerland and the United States of America. (see fig. 26)

Having a similar grade however, doesn't imply that those countries have a similar outcome when it comes to state whether a country is happy or not. Australia for instance has had a 7.28 grade on the World Happiness Report and it is ranked 9th (top), while on the Gallup World Poll it has received 7.45 and it is ranked 29th (middle) and finally on the Instagram Happiness Index where it has had a 7.13 grade and it is ranked 70th (bottom), as seen in figures 26 and 27.

A similar setting can be seeing in the other countries as well. For more information, see appendix P and Q.

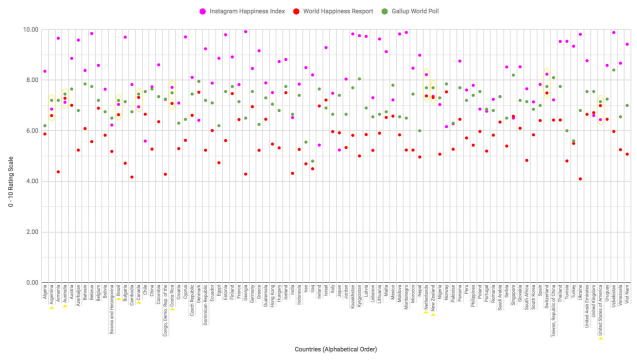

Fig. 26: Scatter Plot graph comparing the data from all three rankings in alphabetical order (Instagram Happiness Index X World Happiness Report X Gallup World Poll).

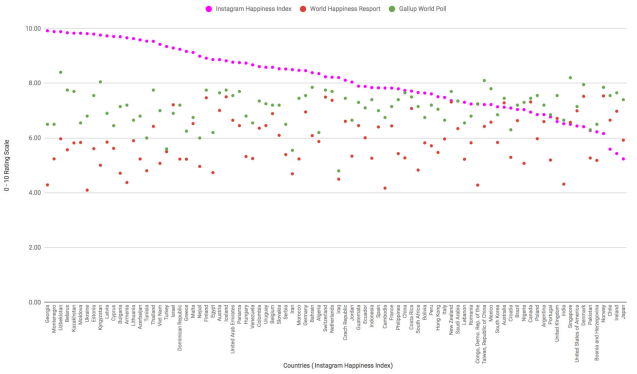

Fig. 27: Scatter Plot graph comparing the data from all three rankings per country - IHI (Instagram Happiness Index X World Happiness Report X Gallup World Poll).

After running the Spearman correlation rank, once again the correlation between the Instagram Happiness Index and the Gallup World Poll was estimated by rho = -0.20, the p value obtained was equal to 0.9505 at the 5% level of significance, indicating that there is an even lower correlation between the rankings, or in other words, contrary to what was expected at first there is no sufficient evidence to suggest there is a significant correlation between the two rankings since they are not even close to be strongly positively correlated. (see table 15) For more information, see appendix N and O.

Spearman's rank correlation rho			
Data Source	**IHI ∩ WHR**	**IHI ∩ GWP**	**WHR ∩ GWP**
S	169540	136512	44929.2
p-value	0.9985	0.9505	0.9636
rho	-0.3064806	-0.2020740	0.6043704

Table 15: Output for the Spearman's Rank Correlation in R statistics environment (Instagram Happiness Index X World Happiness Report X Gallup World Poll).

12. FUTURE WORK

Either running the software for a longer period of time or collecting small databases in order to yield multiple groups of information could be beneficial in this sort of research meaning that it would eventually be possible to compare and analyse different datasets in a much more diverse set of information.

An upgrade that could be made in order to make the data gathered more reliable is to look into the words from the GEW and have a better and more accurate translation from these words into the 11 most spoken languages in the world. Equally, a better choice of words could result in a more assertive and reliable research from a technical point of view. For instance, in our paper we decided to use the word amusement as a noun and not amused as a verb in the past tense that could lead into a much bigger dataset. Another example could be replacing the word joy for happiness, which could be way more popular when thinking of social media hashtags and then it could produce a greater amount of information to be analysed.

Another way of expand the research could be to fix some bugs in the code, as collecting data from the Bengali language was not possible because Instagram kept returning the message 404 i.e. word not found, and because Punjabi and Hindi languages also had the same issues although in a much smaller scenario, having words that could be found while others kept getting the same error message.

Even though the top eleven most spoken languages in the world already covers 47.46% of the list by number of native speakers (table 3), a further review, when it comes to the number of words, would be to increase the list of languages spoken aiming to collect a bigger amount of data i.e. hashtags.

Just like many other social networks it is not mandatory to choose a geotag location for a post neither there is a restriction in choosing the exact location for a place, meaning that any user is allowed to take a picture in Los Angeles but tag the location as if it was taken in New York for instance. Therefore, an update either in the code or in research itself could be made in order to close this gap of misleading information. A good example can be seen in fig. 28, in which #belgium was used in the post, but the location of the picture and the caption are marked as Skive, a small town in Denmark, as well as another caption written #denmark. In addition, the date is inaccurate because the post is dated March 22, 2016.

Like previously mentioned fake Instagram accounts, called bots, are also something to be taken into consideration, nonetheless it is very unlikely that a study could get rid 100% of such robots. Since bots can make a post and then use a random hashtag that in the end has nothing to do with the post whatsoever. Solving both problems above would potentially require an artificial intelligence working as part

Fig. 28: Hashtag #belgium (English) and #belgien (Danish) tagged in a post by a user. From Poulsen, Mia Bergmann. Photo of a candle light. Instagram, 22 Mar. 2016. www.instagram.com/p/BDREZzwjuw0/

of the code aiming to check whether the image from the posts has connections with the hashtags and in a further stage, to check connections with the caption of those posts. By doing so, a much bigger number of bots would be potentially identified and then ruled out from the outcome of the research making the data more accurate.

Another solution for the bots would be to develop a script that estimates how much of the posts are bots, based on a bigger sample of the dataset, thereby an estimation of how much data should be thrown out from the total amount of data could offer a more accurate outcome for this kind of social media research.

An even more complex issue is when we have two countries fighting over a given region, for instance the conflict between Hong Kong and People's Republic of China (Boland), or the conflict between the Russian Federation and the Crimean Peninsula ("Annexation"; The Associated Press and Reuters) and lastly the centenary conflict between Argentina and Great Britain for the Falkland Islands also known as Islas Malvinas in Spanish ("Falkland Islands"; Mingo; Foot).

Finally, it would be interesting to find a solution for cases in which a geotag location is made in areas with missing or unclear data and for that reason it is difficult to pinpoint a specific country for a post.

All aforementioned improvements would make the research not only more valuable but also more reliable, however they could potentially change the outcome of a future paper. It would be prudent then to dive deep in this topic in order to trim off the edges and thus get a better and more accurate result for a greater outcome for the research as a whole.

List of Limitations
1. Running the script for a longer period of time
2. Yielding multiple groups of information - several datasets
3. Improve translation accuracy for all 220 GEW words
4. Improve date accuracy in data collection
5. Fix bugs in the code and/or make improvements in the script
6. Increase the list of languages spoken
7. Identify and exclude bots from the results
8. Fixing missing or unclear geotag location data

Table 16: List of limitations that could possibly be considered for a future work on this topic.

Summing up the paper, the main reason for doing this research was to demonstrate that the Instagram Happiness Index offers a different view when it comes to analysing happiness on social media networks, which seems to be completely distinct from other gold standard studies already know by researchers in this field such as the World Happiness Report and the Gallup World Poll. Our study aims to show how users portray themselves online rather than having an overview about happiness in a given country like other existing studies usually do. Thus, the Instagram Happiness Index could be improved further over time making it a reliable, respected and accurate source of information of how happy people portray themselves on Instagram. In the end, it could eventually become a small part of major studies on happiness by having a share or weight of the results coming from social networks.

13. ACKNOWLEDGEMENTS

I would like to thanks to Peter van der Putten and Ruut Veenhoven for supervising my graduation research.

A special thanks to Bruno Maciel Peres for collaborating in the technical aspect of the project such as programming, data collection and the usage of Instagram API and Code Grid Library.

I could not forget to thank Maarten Lamers, who provided me with great insights and helpful comments in the early stage of this research.

Additionally, I am grateful to all media tech students who in one way or another gave me the courage and stimulus in the project and supported my study.

Lastly, I appreciate the support from my family during all my graduation in the Netherlands at Leiden University.