UNIVERSITY OF GOTHENBURG

DEPARTMENT OF EDUCATION, COMMUNICATION & LEARNING

INFORMATION EXCHANGE IN COMMUNITY QUESTION AND ANSWER SITES

An empirical study of Stack Overflow as Community of Practice

Panagiotis Rafailidis

Thesis:	30 higher education credits
Program and/or course:	International Master's Programme in IT & Learning
Level:	Second Cycle
Semester/year:	Spring term 2020
Supervisor:	Markus Nivala
Examiner:	Mona Lundin
Report no:	VT20-2920-007-PDA699

Abstract

Thesis:	30 higher education credits
Program and/or course:	International Master's Programme in IT & Learning
Level:	Second Cycle
Semester/year:	Spring term 2020
Supervisor:	Markus Nivala
Examiner:	Mona Lundin
Report No:	VT20-2920-007-PDA699
Keywords:	Stack Overflow, Communities of Practice, question, comments, information exchange, CQA

Purpose:	The purpose of this study is firstly to explore how the users of Stack Overflow, an online community-based Question and Answer platform, inquire for information through the formulation of a question, according to the community guidelines. Secondly, to observe the information exchange procedure between the question and the comment section. Finally, to investigate, if this volunteered informal learning and information exchange environment shares any similarities to the Communities of Practice.
Theory:	An emphasis was put to analyse situated learning and Communities of Practice as described by Wenger (1998) as well as additional research dedicated to this particular field. Additionally, previous researches were analysed to observe the important characteristics regarding the question formulation and information exchange (suggestive solutions).
Method:	Two different mixed method content analysis were executed. Firstly, to define the patterns that entwine questions in the platform and observe the importance of the community's guidelines. Secondly, another mixed content analysis to observe the interaction between the asker and other users in the comment section. The data was gathered as part of a research project between the *Faculty of Education* and the Faculty of *Applied Information Technology (IT)* within the University of Gothenburg.
Results:	The first part of the analysis showed that the community guidelines are the epitome for a successful question in the community. Even though users are willing to aid new user with ill-informed questions, there is a connection between the question's score and asker's reputation. As thus, users with longer expertise in the platform are generally able to create successful questions. The users of the platform are utilizing the comment section as a troubleshoot chat; aiding the user not only by providing suggestive solution but by reformulating their question for future use by other users. Users of the platform edit and provide external resources to learning material to aid askers that are still learning how to code. It was evident that they do not recognise the platform as a hand-out solution site but as a community willing to aid users that are willing to learn.

Foreword

I would wholeheartedly like to thank my family for making it possible to finish my studies through their invaluable support. Additionally, I would like to thank the close friends and classmates that aided and encouraged me while I was working on my thesis.

Special thanks to my supervisor Markus Nivala for his valuable feedback, Senior lecturer Thomas Hillman for his ideas during my internship, as well as the PhD students Svea Kiesewetter and Alena Seredko for their helpful hints and feedback.

Table of content

List of Abbreviations

Abbreviation	Explanation
MOOC(s)	Massive Open Online Course(s)
Q&A	Question and Answer sites
CQA(s)	Community based Question and Answer site(s)
SQA(s)	Social Question and Answer site(s)
SO	Stack Overflow
CoP(s)	Community(ies) of practice
RQ	Research Question
API	Application Programming Interface
HRQ	High Rated Questions
LRQ	Low Rated Questions
HRUQ	High Rated Unanswered Questions
N	Refers to Total Number
FAQ	Frequently Asked Questions

1. Introduction

Programming is a never ending journey of knowledge. As with any other profession that requires long-life learning, programmers are presented with versatile way to learn during their career ranging from pure formal environments (classroom education), to more digitalized solutions (such as Massive Open Online Courses and webinars) as well as informal learning environments. Informal learning theory has been valuable for the evaluation of learning available to adults as they mostly learn outside of formal education contexts (Gray, 2005; Marsick & Watkins, 2001). Programming in that regard has been considered important not only as a professional tool but also for the development of the individual learner through an increase in communication and cooperation skills (Sonnentag, Niessen, & Volmer, 2006). Even though previous research tended to view informal learning as an individual process (Eraut*, 2004; Gray, 2005; Ziegler, Paulus, & Woodside, 2014), given the aforementioned attributes, researchers have moved the focus from individuality to group-setting informal environments. Even though informal learning for professionals has been evident outside the corporate world, usually previous research has focused on informal company settings (Johnson, Blackman, & Buick, 2018; Manuti, Pastore, Scardigno, Giancaspro, & Morciano, 2015; Marsick & Watkins, 2001). One of the latest shifts in *social learning* for professionals has been seen in the form of online communities where users tend to engage in conversations and focus around a common interest, while sharing knowledge to community members through their shared repertoire (Daniel, O'Brien, & Sarkar, 2007; Gray, 2005). Conversations between users are used as the information exchange tool and through the observation of those, researchers are able to investigate *learning as it is happening* (Ziegler et al., 2014). Professional programmers are usually engaging in similar communities for their development, which are in the form of forums or online social Question & Answer sites. An example of the latter is Stack Overflow. According to the numbers collected from Stack Overflow (SO), more than 100+ million programmers are currently using the platform creating thousands of posts, and answering questions. The platform of SO has become a prominent webpage for a big percentage of developers, with sufficient English skills worldwide, who turn to SO for the acquisition of new information through its peer interaction.

Programmers that utilize the platform are able to create smaller communities filled with information, based on their previous experiences and solutions to issues that they have faced. Those smaller communities are based on the interest and expertise of the users (e.g. a Python language community). In addition to the creation of a question-answer thread, users have the ability to comment, edit, flag, close as well as upvote/downvote the information shared (Sin, Lee, & Theng, 2016). When commenting on existing questions, users discuss, provide solutions and aid the asker to achieve the closure of the particular thread. This *shared repertoire* is judged by the voting system by the rest of the community and sets the standards as to what questions, answers comments, users of this community find well-structured against material that does not follow the SO guidelines. The way that knowledge is shared in SO and other similar community-based platforms, possibly resembles the practices of online communities of practice and situated learning, as described by Rosenbaum and Shachaf (2010), where users learn through the experience of others through *social interaction* and *volunteered contribution*. While research about the quality of question/answers as well technical features around SO platform is extensive, research focusing on SO platforms as learning environments is scarce. Specifically, the research seems quite limited around the connection of this particular community with learning theories such as experiential learning and situated learning as described through the communities of practice (CoP). Through the extrinsic motivation that the site provides can encourage users' participation, it does not indicate that the users are joining these Q&A sites with the goal to learn or if they see the site as a hand-out answer book.

In this particular study I will investigate: 1) the distinct characteristics regarding the formulation of questions being posted by users in the SO platform, according to the score rating system (upvote/downvote), 2) the discussions developed around the formulation of the posted problem and its potential solution provided by the community, in the comment section of the same questions.

2. Background

In this section, a brief overview will be provided regarding the transition from offline and general Question and Answer sites to what is currently known as social or Community-based Question and Answer (CQAs) sites. Through already conducted research, the main features of CQAs will be provided, along with an overview regarding the users and their participation in the information exchange.

2.1 Before CQAs

Long before the internet was available to everyone, people were able to obtain information and interact with each other through offline mailing lists and informal personal groups. Therefore, except from the knowledge gained through the instructor-student interaction, in Skinner's behaviouristic classrooms (Skinner**,** 1968) people seemed eager to attain knowledge from the equal users of the community through social interaction (Vygotsky, 1964). According to Sowe, Stamelos, and Angelis (2008)**,** with the early rise of the internet, online mailing lists were a common medium that users used to communicate and exchange information. Knowledge seekers and knowledge providers were able to exist in the same environment, interchanging roles and create a community of practice where each individual learns from their community.

While more and more individuals gained access to online mailing lists, forums, etc. through the widespread of the internet, many companies saw the opportunity to evolve this need for social interaction and create something new. This new format of interaction was Question and Answer sites which would shape the information exchange in the future. This transition would attract many users in comparison to, up until then, established methods. According to the research of Vasilescu, Serebrenik, Devanbu, and Filkov (2014), when they analysed the activity of r-help and similar stack exchange sites, they observed that the questions asked in the latter were ever increasing in substantial rates.

One of the first versions of Q&A sites, were created as Professional-guided, paid expert-based Q&A sites. Simply put, a team of professionals around a particular field were responsible for providing answers to the users that contributed with questions. Thus, the information exchange usually did not take place between users of the community but through the interaction of users with an employee. One example of these were Google Answers, where users were able to ask questions around a subject and obtain information through a plethora of employees that specialized in this particular field (Regner, 2014)**.**

At the same time the first Community based Question and Answer sites (CQAs), the predecessors of today's famous CQAs such as Reddit, SO, Quora, appeared in the form of forums. The difference with mailing lists was that users were finally able to have a more immediate access to information through the massive amount of users that those sites were able to collect (Chua & Banerjee, 2015). Instead of relying on the willingness of an employee, now they had the chance to aid and be aided through interaction with other users. The success of these sites was related to the level of activity from the users. The participation and interaction were considered higher in CQAs in comparison to traditional mailing lists, while the knowledge providers reacted significantly faster (Vasilescu et al., 2014).

2.2 Community Q&A sites (CQAs)

CQAs are famous for allowing any individual to post questions at any given time, while having other users answering those questions (Shah, Oh, & Oh, 2008). One benefit of any Q&A site is none other than the interaction between the users, which have been considered important for an effective learning environment (Chao, Hwu, & Chang, 2011). The user asks a question about a real-life problem that they are facing, as they need somebody to help them understand the solution to their problem (Seaman, 2002). The answerer could be any user of the group while most of the times the users are using their natural language to ask questions, obtaining personalized answers which seem to be the preferred way of answering for these professionals (Plass, Moreno, & Brünken, 2010; Ponti, 2015).

The content of discussion in Q&A sites could greatly vary.

According to Harper, Moy, and Konstan (2009) questions could be categorized in two groups:

- informational questions: questions asked with the intent of getting information that the asker seeks to receive (problem solutions)
- conversational questions: questions asked with the intent of stimulating discussion (comparing languages, different solutions, etc.)

Another categorization has been done for the answerer. According to Gazan (2006), answerers belong in two types: specialists and synthesists. The first one provides answers based on their existing knowledge without external resources, unless it is to support their argument, while the latter provides answers using external sources without claiming any expertise, most known Q&A sites such as Quora.

Where most CQ&A differ though is in their structure and design. Some of these try to create a completely anonymous environment where everyone is free to ask and provide answers in a plethora of topics (Reddit). Others allow the users to create a more personalized profile which can be customized around their interests as in any other social media platform such as Quora, a similar CQA site.

2.2.1 Answer credibility factors

Many users browse through the threads of CQAs each day looking for solutions to their problems. A percentage of those have no particular interest to become active users of the community, memorize the correct do's and don'ts of this social environment while at the same time remain skeptical about the validity of the answers provided in these forums.

The most valid critique that Q&A sites usually get, is based on the credibility of the information. The reason being that anyone can post an answer, without a peer-review process, which can lead from a well-established answer, to abusive/spam answers (Su, Pavlov, Chow, & Baker, 2007). The credibility factor is based on the believability, trust, accuracy and objectivity, of the answer, among others (Self, 1996). Potentially this leads many users to sharpen their skills through a MOOC program or other digitalized formal forms of learning. The information exchange in CQA sites is highly dependable to the users' skills, literacies and intrinsic motivation to learn. As a result, the user must critically make a credibility judgment which according to Fogg (2003) and his interpretation theory, it is a two-stage process. Firstly, the user has to observe the elements that entwine the website and later based on their observation, make their interpretation. These elements can range from the tools, material and anything observable from the user which could impact their judgement.

3. The empirical context of the study: Stack Overflow

In this section a deeper look into the SO and its design will be presented. An analysis regarding SO features, from general to specific, was conducted, along with the inclusion of literature to investigate how its design shapes the interaction and user engagement.

The creation of SO is just a part from a larger network, created in 2008 which is known as Stack Exchange. Stack Exchange is a family of CQA sites (more than 100sites), varying in topic and fields of expertise from programming to English literature, Physics, Sales, etc., with most famous ones being the SO, Super User and Ask Ubuntu (StackExchange.com).

One common theme among most of them is that the creation of information is solely based on the community and it is rewarded with reputation points based on Stack's reputation system. The users are able to ask questions and use tags to attract other users with the same expertise and interests.

Specifically, SO is a question and answer site for professional and enthusiast programmers. Professionals can interact with each other, acquire answers on their own questions, while they can also browse in a forum based library with an unlimited amount of information available from previous questions and answers (Vasilescu et al., 2014). Different types of users are attracted by SO, with some of them being professionals as well as pure hobbyists. Stack receives around 8000 questions per day (Meta.Stackoverflow.com) and until 07/09/2019 the platform had received more than 18 million questions. Questions posted in SO can remain open as long as they comply to the platform's guidelines. This has as a result that scores and user's reputation can change considerably in a year's time. At the same time questions can potentially be closed, edited and commented even if an answer has already been received (see Results section/reference).

3.1 Design of SO

3.1.1 Reputation System

A person that attends a CQA is usually guided by the intrinsic motivation to acquire information. Albeit, intrinsic motivation might be just enough for the asker, one could argue for the answerer's willingness. According to literature: the answerer could be 1) guided by the intrinsic passion to learn about the problems of the community, 2) want to provide with their insights or 3) in the thought that by providing aid in this particular time, someone in the future would do the same for them (Lakhani & Von Hippel, 2004; Vasilescu et al., 2014). Possibly few users of SO try to answer questions based on the feeling of cooperation and mutual benefit but that is not always enough. For this reason, SO creators, composed a rating system instead.

By participating in different activities, - questions, answers, edits, etc.- users of SO can increase their reputation score which gives more capabilities to the users. Each user begins with the reputation score of "1" and through different activities they are able to increase their reputation in different increments (by 1, by 5, etc.) depending on how important their action was. This plethora of increments somewhat explains how a part of the community is able to achieve ratings higher than 300.000 or more. This rating is always visible to the other users, providing visual feedback as to the knowledge of the answerer (see Figure 1). Through visual rewards in the format of a reputation score, medals, and other achievements the user is presented with extrinsic motivators to create a come-back relationship between user and website (Mamykina, Manoim, Mittal, Hripcsak, & Hartmann, 2011; Ponti, 2015).

Even though the reputation system plays an important role in the engagement of the users, no system is perfect. There are examples where gamification, or other extrinsic features, are too much that result in a negative effect. In these cases users are potentially losing the intrinsic motivation to participate in

the activities as their main goal is to reach the end game exploiting the gamification features for personal gain (Hsieh, Kraut, & Hudson, 2010). According to the survey of (Calefato, Lanubile, & Novielli, 2018), many users tend to agree that SO favours the "rich get richer" effect. As a result, the users of the community tend to prefer answers given by an expert user of the platform (with higher reputation), as by judging by their level of reputation the answer yields bigger significance than the one provided by a new user. According to Vassileva (2008): "SO has three properties of new social learning technologies: support learners to find the content that they seek, the ability to connect with the right people as well as motivate people to learn by the inclusion of its reputation system*"*. Even though the two first ones are observable based on the statistics provided by SO and the opinion of the users, the third one is still under debate.

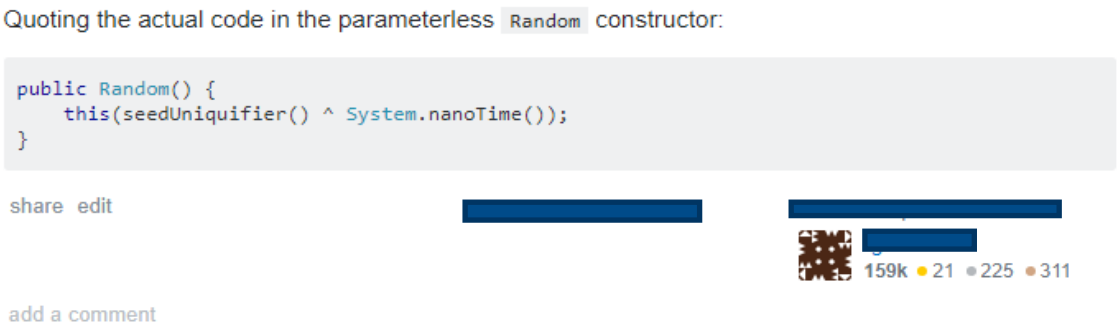

Figure 1: User's Name with the reputation level (159K), along with the medals that they already own.

3.1.2 Medals

Similarly to the overall rating, users can also gain achievements in the form of medals to mark certain accomplishments, which are later divided in bronze, silver, gold, based on the level of difficulty (see Figure 2). Accordingly, those medals promote different activities that are important for the interaction of the community and depending on the difficulty of the task the medal shapes from bronze to gold. So far, we can see that the creators carefully used, what could be described as gamification features (Kafai & Burke, 2015) to attract the users, provide an extrinsic boost in order to maintain a healthy, come-back type social interaction (see Figure 2).

As a result, it has been seen that users' behavior is directly affected by the badges and possibly similar functions provided by SO, as "users (i) are rewarded with points to encourage the desired behaviour (and may be subtracted points to sanction undesired behaviour); (ii) are awarded badges after collecting sufficiently many points or when performing certain activities; and (iii) have their progress tracked and their achievements displayed publicly in a leaderboard, to create competition between them" (Vasilescu et al., 2014, p. 3).

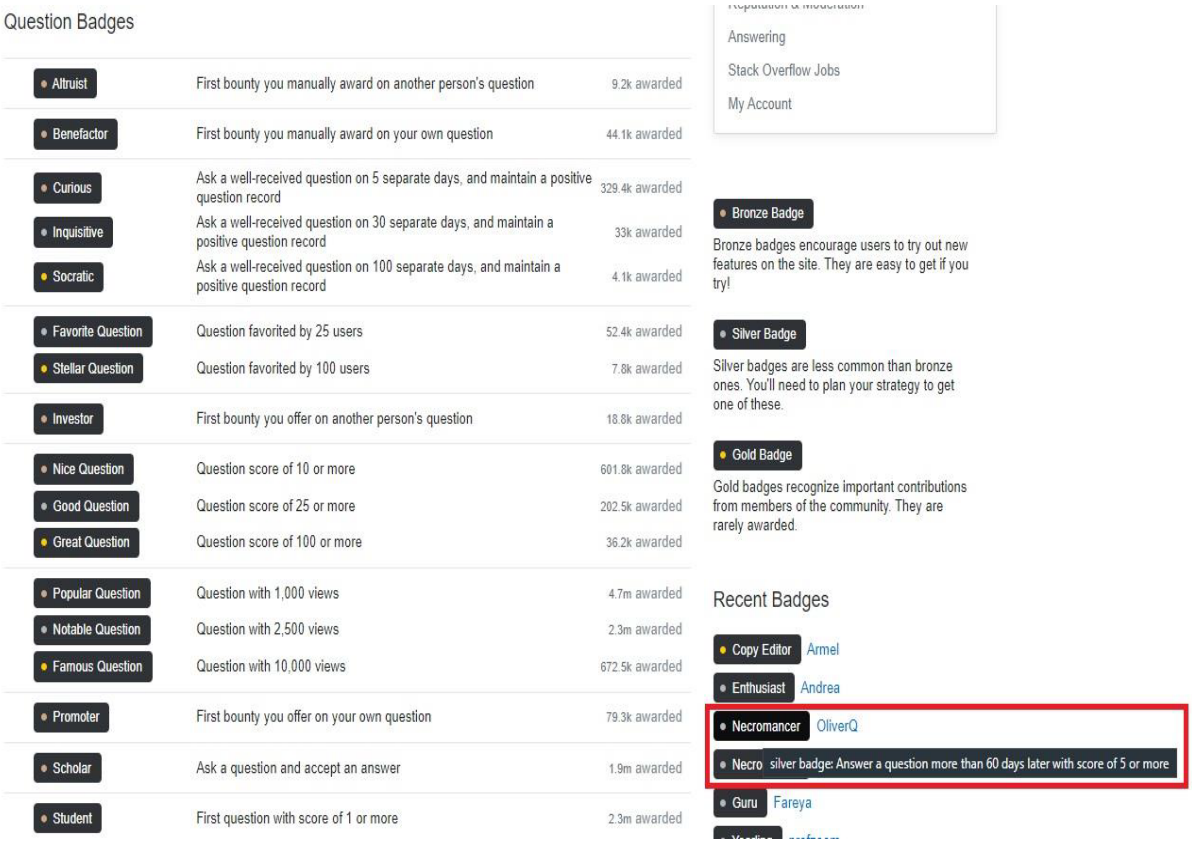

Figure 2: A collection of SO badges. Marked is a badge to urge users to answer old unanswered questions.

3.1.3 Voting System

Being an open site, filled with an abundance of information, a user might struggle to correlate the existing answers to their need All the users participating in the question-answer process are able to upvote or downvote each given information, making it easier for the asker as well as the users to find a suitable answer. It should be noted that not all users have the chance to affect the up/down vote process. Users with reputation less than 15 -a reputation score which is considered quite low and can be achieved with few contributions- even though they can physically select the up/down vote button, it does not make any changes in the final score

3.2 Question Creation Guide

Stack has clear guidelines when a user creates a new question. There are 3 main steps that a question must follow:

1. Summarize their problem with the inclusion of details about their goal, describe expected and actual result with the inclusion of any error messages (Explanatory Text).

2. Describe what they have tried so far, along with the inclusion, if possible, of any information found on the SO site (previous posted questions) or any other external link and why it didn't meet their needs.

3. Provide some code snippet, so that others can observe the issue at first hand and replicate the code, if needed.

A descriptive guide is provided to the user, around the question body organization, emphasizing the importance of the first paragraph, which should introduce the problem, starting from a general information-rich description to a more specific analysis of the issue at hand. It is then suggested to the user to include specific code snippets that are related to the issue, if needed, with the avoidance of screenshots of code, data etc.

Finally, the user is requested to proof read the questions and respond to any feedback provided in the comment section by other users. They should argument and provide feedback around the answers collected and be ready to edit the existing question, providing additional information if requested.

More specific guidelines are provided regarding the creation of the title along with the inclusion of tags, among others:

3.2.1 Title

It is advised to the user to be specific with the creation of the title, imagining they present the question to another person. It is distinctly mentioned that a title must be interesting, to attract the users' attention who would later go on with the asker's question body. Even though it is not specifically mentioned as to if the format of a heading or question is preferred, SO provides some examples as to what it is considered a bad or a good question. By having a closer look to these examples, it is evident that all the bad titles are in the format of a title (heading) while the good ones in the format of the question. Though it should be noted, that the provided bad examples are lacking any consistency, information, and syntax so the format of choice could be unrelated to the creation of a successful title.

3.2.2 Tags

The platform suggests that tags can aid the user to attract the right people that can answer their questions. It is advised that the user should provide up to 5 tags to briefly describe what the question is about. Furthermore, it is suggested to start with general tags, crucial around their questions and include specific language numbers (e.g. Python 3.1), if needed. The usage of popular tags is advised with the choice of creating new ones either by the user or by the community if they do not find the perfect tag for their question.

3.2.3 Edits

Sometimes users tend to hurry and provide an ill-informed question. Another feature available in SO, is the ability of the original asker or high reputation users to edit an already posted question. These changes can be related to the formulation of the question, based on the feedback gathered in the comment section beneath the questions. The platform is strict on the correct usage of the code snippet function. The asker should not post the entire program in the question body but only the necessary parts. Correspondingly, it should be mentioned that the moderators are trying to keep the discussion civilized by deleting harassment comments, or solutions that tend to not be part of the problem. Another usage of edits, most usual from users that have enough reputation points to utilize it, is ***bounties.*** According to Stack, if the user thinks that they have created a well-formulated question and they are still not receiving answers, they could draw more users for potential answers through the utilization of a bounty which features the question in the homepage featured tab. In order for the user to create a bounty they have to "sacrifice" some of their reputation points (in increments of 50) which are rewarded to the answerer and are not refundable. Even though bounties can be set at any given time, users tend to utilize them mostly once the question stays unanswered for a while which entwines bounties as part of the "Edits".

3.2.4 Comments and Answers

Answering a question is not the only way for interaction with the community in a SO thread. It is most common, that the utilization of the answer function to be a mean to solve the asker's issue. In comparison to other CQA platforms as Reddit, answers are not commonly utilized to create meaningful discussions or argumentations around the formulation of question or the issue at hand. Stack suggests that those forms of discussions should take place in the format of comments. The creation of comments is available in two distinct sections on the SO platform. One section dedicated beneath the questions, and separate sections beneath each answer provided. The SO guideline suggests that users should provide feedback around the question in the comment section beneath the question, while the asker, along with any other user, will provide feedback on the solution in the comment section beneath each answer.

This feature is not available to everyone as, through Stack's gamification features, a user must reach a reputation level of 50 before being able to comment on another person's question. Even though research has already been done as to which criteria formulate a good answer, only a handful of researchers have focused on the interaction in the comments. One hypothesis around the limited research around comments, could be the reason that comments do not generate any reputation points. According to the research of Zhang, Wang, Chen, and Hassan (2019), there were occurrences where the comments tended to have more up-votes than the selected best answer, proposing usually corrections or re-formulations on the selected best answer or initial question. This can happen in the form of a snippet (code) or through additional explanatory text. Even though SO sets the comments in a subsidiary position, users are utilizing this feature in many posts, setting their disbeliefs and corrections, while prompting the reputable users with the editing skills, to reformulate answers and questions, based on their corrections suggested by the rest of the community.

Summing up, SO is a platform that provides programmers a place to interact and share information. In order for the information exchange to be fruitful, SO provides a guideline around the creation of a question as well as extrinsic motivating factors to boost user's (answerers') engagement. This information exchange is achieved through the creation of threads where users can engage in the creation of a solution regarding the posted question but can also provide feedback through a plethora of features such as voting questions, comments, editing, flagging, etc.

4. Previous Research

Having introduced SO as a CQA website, I will now present previous research in relation to a) how professional programmers use questions for the acquisition of information and (b) how this interaction (mostly visible in the comment section) between the users on social platforms as SO shapes throughout the active users of the platform.

Previous studies have focused on how the users interact in CQ&As while a substantial margin of those researchers has focused on SO as to what characteristics usually entwine a good question or answer based on the users' rating. This chapter will be divided in two smaller parts, with focus on the formulation of the questions and comments/answers in CQAs.

4.1 Formulation of Questions

According to the research of Allamanis and Sutton (2013): "Question types represent the kind of information requested in a way that is orthogonal to any particular technology." Questions could be divided in categories based on the focus of the asker. Some of them aim to gain solutions on real-life issues, usually having code snippets in the focus, which resembles the idea of a "worked example" (Plass et al., 2010). Other questions focus on the learning part of, example given, a new programming language and are usually filled with external informative references (Harper, Moy, & Konstan, 2009). Of course, both of these characteristics can coexist in a single topic. The paper of Allamanis and Sutton (2013), through the utilization of Latent Dirichlet Allocation (LDA) a generative model for describing documents based on co-occurring words, showed that the types of questions are similar, independent of the programming language, while some of the main topics that users focus on were 1) *concept questions*, 2) *requiring solution (*when something is not working), 3) requesting aid in learning a new language, among others.

The answerability of a question directly affects the chances for it to be considered successful by the other users of the community. In the platform of SO a number of questions could possibly stay unanswered, receive a negative score or never receive the answer that suits the asker's needs (Calefato et al., 2018). As a result, the way that the question is formulated plays a significant role in the answerability of the question. A well thought and structured question is more prone to receive multiple user answers in comparison to an unstructured and ill-informed question (Li, Jin, Lyu, King, & Mak, 2012). According to (Chua & Banerjee, 2015) and their literature review, there were three distinct metadata features that could contribute in the creation of an answered topic. The first was the asking time-window of the question, as it was also suggested by the previous study. According to Allamanis and Sutton (2013) it is evident that weekends are less busy in certain languages which focus mostly on corporate environments (SQL Server), while other languages more common for enthusiasts such as Python and C++ are prominent during the last days of the week. The second was the reputation score of the user, referred as *recognition (Chua & Banerjee, 2015).* The third being the *popularity* of the asker. A plethora of Q&A sites or forums provides the opportunity to the users to be divided in even smaller communities, structured around their interests. This has a result that particular users that contribute frequently to be recognised by the rest of the community. Subsequently, the *popularity* refers to the level of recognition of the asker from the rest of the community, which does not always correlate with the reputation score. Furthermore, the popularity has been described as *derived,* as the users have the ability to up-vote and down-vote the activities of each user, thus affecting the "acceptance" of a question along with their reputation score (Chua & Banerjee, 2015). According to the findings of the study, questions with numerous down-votes and short duration of membership were most likely to be unanswered. Additionally, another finding in the same study showed that, probably out of altruism, high-reputable users tend to help new users, while many times, the level or popularity or reputation seemed irrelevant to the answerability of the question. This contradicts previous studies

that showed that the level of reputation of the user played a significant role in the acceptance of a question by the community (Yang et al., 2011)**.**

According to Calefato et al. (2018) and their conceptual framework regarding factors of influence, even though the chance increase for a question to be considered successful regarding the users reputations was low but statistically significant, the difference of expertise between a new user and a low reputable user was high. Thus, the higher the expertise of the user, the more chances for a question to be created according to the community's guidelines. Previous studies have also shown, with the creation of a conceptual framework and empirical analysis, that questions with shorter title and description as well as fewer tags tended to attract more answers (Chua & Banerjee, 2015). This verifies the research of (Calefato et al., 2018) who have suggested that a concise writing style would increase the probability of success, even more if the text was followed by a snippet of code, but at the same time contradicts their results around the importance of a *shorter title.* Furthermore, the idea around tags comes against the effort of (Saha, Saha, & Schneider, 2013) and their tag suggestion model for filling tags for the askers which was analysed before. Tags are an important feature that with its correct utilization, can affect how many users are attracted to the particular question, affecting its success rate. According to (Saha et al., 2013), not all users utilize the five available tags to their questions which has a result many questions to remain unanswered or attract the wrong users. Even though, the users are presented with a "manual" as to how to formulate their question, provide code, etc., it is hard for a new user to follow it consistently. One suggestion made by the pre-mentioned researchers has been an automated tag system which would firstly successfully assign tags to existing questions while providing new users with suggestions as to which tags could attract the best answer for the user while warning them about probable incorrect tags (Saha et al., 2013). It is possible for new users to get down-votes for their questions with no feedback as to what they executed wrong, resulting in "ghost users" that passively observe posts without having the courage to formulate questions. According to the case study of Srba and Bielikova (2016), during 2014 the number of questions that did not receive a "best answer" or that got deleted (because of community violations or remained unanswered) exceeded the number of questions that show their questioner's need fulfilled.

The Stack guidelines, along with the community, tended to agree that the "writing tone" of the user for formulating a question plays a significant role in the up-vote/down-vote system. More specifically the users are asked to keep the conversation formal without signs of negative or positive sentiment as that could hinder the question's success (Calefato et al., 2018). More often, the new users tend to express negative sentiment towards their self or show "gratitude" to the users contributing to their issue. While the reputation increases it has been observed that this phenomenon reduces.

4.2 Formulation of answers and comments

A well formulated question does not only provide the asker with more reputation points and could potentially lead to an up-voted question but it also sets the standard as to which users it will attract affecting the quality of the answers and comments that it will receive. According to SO guidelines users should avoid questions that are too broad or could be considered opinion based. Previous studies have showed that, complex questions that lacked clarity could have potentially remained unanswered or have attracted less users and this seemed to be a common pattern among even more conversational CQAs (e.g. Yahoo Answers!) (Asaduzzaman, Mashiyat, Roy, & Schneider, 2013; Chua & Banerjee, 2015; Yang et al., 2011). In comparison to other SQAs, SO provides the opportunity to the users to comment and edit questions and answers separately. As thus, the utilization of answers in SO is predominantly used for the provision of a successful solution to the issue posted by the asker. Fewer are the examples where answers are used to provide feedback to the initial code and question body of the asker.

Accordingly, SO provides 2 distinct categories for comments: 1) Comments beneath the question body to aid users to view the comment in context and 2) comments beneath the answer body for clarifications, updates and other suggestive information to the corresponding answer (Sin et al., 2016). As a result, while a question might have received multiple answers, on average each question received 2,04 comments (Sin et al., 2016). From the comments/edits provided, most of the collaborators has been shown to be done by users who do not possess an answer related badge (Adaji & Vassileva, 2016). It thus has been suggested for SO to incorporate strategies to encourage this active collaboration through edits and comments as it does for the formulation of questions and answers (Adaji & Vassileva, 2016; Soni & Nadi, 2019).

The research of Sin et al. (2016) , through a social sequence analysis (SSA) showed that askers that engaged in multiple comments with other users would result in better outcomes for the asker, which could be resulted in a higher question score or receiving more answers. According to the three-step heuristic analysis of the SOTorrent database, Soni and Nadi (2019) observed that questions with more upvotes were likely to attract a bigger percentage of users willing to provide answers. Similarly, being able to create a question that attracts many users who collaborate in the comment section can have a different nature of contribution in relation to the provision of answers such as more diverse perspectives and recommendations (Adaji & Vassileva, 2016; Soni & Nadi, 2019).Similarly comments, except for the users' interaction and discussion, could provide significant changes to flawed posted answers (Soni & Nadi, 2019). The same could be argued for the aid they provide in the formulation of questions. The relationship between comments and updates is not that clear. The same study, showed that through an extensive analysis across five languages, only 4,6% of comments resulted in an answer to be updated. Another percentage (8,7%) showed that comments could contain discussion-focused text with no improvements over the corresponding answer. Even if they were aiding in the reformulation of the answer, 27,5% of those did not attract the attention of the answerer.

Though a logistic regression framework for the analysis of the actionable factors, users of SO seem to agree with the guidelines of the platform as to what is considered a well-established answer (Calefato, Lanubile, Marasciulo, & Novielli, 2015). In a qualitative analysis to investigate what makes a good code example of Nasehi, Sillito, Maurer, and Burns (2012), users seemed to aid the questioner regarding the formulation of their question by providing answers that re-formulated the question in a more structurally correct manner. Sometimes this restructure would result in the provision of the best answer attribute as it was enough to solve the asker's problem. This finding is closely related to the deductive coding analysis of Chua and Banerjee (2013) where the formulation of a question in similar CQAs would attract different kind of answers, where a "factoid" question, would result in a higher percentage of answers. Additionally, the same study showed that few answerers were willing to provide new ways of improving the code's readability and efficacy by suggesting alternative "routes" to the asker's issue, even if that was not the initial intention of the asker. Moreover, users tend to divide the initial question in smaller pieces, and give explanatory answers for each part of the code.

This, in addition with qualitative case study around the importance regarding the utilization of comments in SO (Zhang et al., 2019), could possibly show that online communities of practice, through collaborative learning, establish new norms to achieve the end goals (learning outcome) while fostering a sense of co-accomplishment, through the repetition or re-evaluation of the question and answer process (Berlanga et al., 2008; Lee, 2004). The gamification features of SO play a significant role to this collaboration as researchers have recognised the importance of features like these (Deterding, Sicart, Nacke, O'Hara, & Dixon, 2011), who set the foundations of the "come-back" relationship of the users and the further attraction for many professionals to the platform. Studies have shown that the inclusion of badges can potentially increase the user's participation (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2013; Deterding, 2012). User's behavior seems to be affected by the badges/medals as according to Wang, Chen, and Hassan (2018) as users seemed to do more revisions (edits) during the days that the system rewarded special medals for this action in comparison to "normal" days. Furthermore, they showed that users who were rewarded with a medal for this kind of action were 17x more likely to perform it again than users without revision-related medals.

Another point that seems significant for the community of SO is the characteristic of providing external references to their answers. In similar forums and professional CQAs, questions are usually related to information seeking (Singh, Twidale, & Nichols, 2009). Those external references can been seen as additional material for *central processing of information or hints of peripheral processing* (Freeman & Spyridakis, 2004). Users of SO tend to provide references in external websites and other learning material even though this might seem to negate the chance of providing a concise, personally tailored to the questioner, answer (Nasehi et al., 2012).

To conclude, according to an interview-based previous research, the answer's tone could negatively affect their reputations among the users of the conversation even if they had high expertise in the field (Kim, 2010). On the other hand, as in any other CQA with a similar system, the higher the rating the more credibility (reliability) they earn (Ponti, 2015) which leads to more chances to have their answer rewarded as the best answer. In both cases of a formulated answer (high rated/low rated), few grammatical mistakes could be overseen, as long as the answer is sufficient (Kim, 2010). Even though, studies have shown that new users could have created their questions in the same level as more experienced askers (Chua & Banerjee, 2015), it has not been thoroughly studied as to if the platform, through the up-voting and down-voting system appreciates the questions and answers of new users in the same level as the pioneers of the field. As in any other community of practice, users need a period to adapt to the new community, learn the formalities, the language and forms of expressions, among others (Lave & Wenger, 1991). This system can be used by the community, to either appraise or disapprove the formulation of questions and answers of new established users. This mechanism is questionable to the amount of feedback it provides to the user as they can potentially be left stranded with no further clarification as to what they did wrong and how to fix it in a future post.

For this reason, the main goal of this study is to analyse what are the characteristics of a high-ranked question for the Python language community and if the comment section can be utilized to provide constructive feedback to the users. A two-step mixed content analysis on: 1) upvoted, downvoted and unanswered questions, 2) the comment section beneath those questions, will address the following research questions:

What defines a high-ranked question, regarding the community guidelines, in relation to unanswered and ill-formulated questions?

How do the users, of different levels of contribution (reputation) interact and inform the exchange of information (asked questions) through the discussion/argumentation in the comment section?

Does (this procedure of) information exchange share any similarities to the characteristics that are evident in a Communities of Practice?

5. Theoretical framework

In this section I will briefly outline the theoretical background of this study, regarding Communities of Practice and their characteristics as described by Lave and Wenger (1998) along with previously conducted research around the information exchange and motivational factors for a successful information exchange community. This analysis of CoPs will provide some insights regarding the similarities to the SO platform, as well as set the basic idea for the creation of the analytical framework regarding the investigation around the information exchange characteristics of the platform. Through the overview of the characteristics that entwine a CoP it will be possible later on to observe if SO have similar attributes through the information exchange of its users.

Information exchange and motivational factors

While corporate environments constantly examine the usage of online networks as a mean to further engage the employees to share knowledge for their further professional development research seems to agree that informal environments can offer the significant potential for knowledge to be shared (Gray, 2005; Hsu & Lin, 2008). Users who participate in those virtual communities, can acquire knowledge through the information exchange by sharing their ideas and thoughts (Chen & Hung, 2010). The platform which enables the form of those communities, usually offers the mediums for *negotiating meaning*, *of making sense of and understanding their work* (Gray, 2005; Thompson, 2011). Researchers though argue that collaborative learning is much more than simply exchanging information (Matschke, Moskaliuk, Bokhorst, Schümmer, & Cress, 2014). Knowledge is treated as a "source of value", emerging from the users of the platform while the environment offers ways for the reuse of the contributed knowledge (Ba, Stallaert, & Whinston, 2001; Gang & Ravichandran, 2014; Matschke et al., 2014; Zheng, Zhao, & Stylianou, 2013). In order for the participants to be willing to share the precious knowledge they have gained throughout the years (Hsu, Ju, Yen, & Chang, 2007) to generate specific domain knowledge (Hsu et al., 2007; Lee & Cole, 2003), the understanding and expectations of an online community are vital parts for the users' engagement and motivation. Expectation is considered a personal trait and research has shown that can vary from person to person as to what a community should offer regarding the connection to other users (Thompson, 2011).As thus, a virtual community platform should provide motivational factors in order to secure the constant distribution of information between the users.

Users, of any Virtual Community, have the ability to either consume information through browsing and reading posts or to provide information by replying in messages and questions (Park, Konana, Gu, Leung, & Chung, 2010).Users tend to be hesitant in sharing information if that sharing does not yield tangible benefits or rewards (Gang & Ravichandran, 2014).According to the social exchange theory(Emerson, 1976), people engage in social interactions with the hope that they will be rewarded in the future when they will be requiring help. At the same time Social Cognitive theory, suggests that a person's behavior is shaped by contextual factors and the person's cognition, as thus the personal actions of the user in a social environment are affected if those actions have personal cognition (Chen & Hung, 2010). Interpersonal trust has been considered by previous research one of the fundamental factors that positively influence users to exchange information (Chen & Hung, 2010; Hsu et al., 2007) (Gang & Ravichandran, 2014) even more when the environment offers the possibility for identification-based trust increasing the familiarity between users (Hsu et al., 2007)**.** If the social environment (online platform) fails to provide users with the ability to build-up trust, users could potentially not share their past experiences in fear of criticism or misleading others (Ardichvili, Page, & Wentling, 2003). The second most met factor was self-efficacy. Previous studies seem to suggest that the sharing and receiving knowledge actions were positively related to knowledge utilization (Chen & Hung, 2010). The platform in which the knowledge is distributed, the ways of evaluation of the quality of the content (Matschke et al., 2014) and the actions offered for knowledge (Chen, 2007) seem to be the third most evident factor for the success of the information exchange.

Learning through experience and cooperation

The forefather of Social-Cultural learning theory, Lev Vygotsky, was the Russian psychologist, who would set the ground principles behind cooperative learning and information exchange, examples of which can been seen in social information-exchange platforms. According to Vygotsky (1964), the idea behind this social perspective is that the community plays the central role in learning, where, through the interpersonal interaction, learning becomes personalized and tends to "make meaning". A fragment of the social learning theory evolved into what is known today as *experiential learning*. The simplest way to describe Kolb's experiential learning theory (Kolb, 1984), is that individuals create knowledge and meaning by sharing their previous, real-life experiences with others.

A theory closely related to the sociocultural perspective of Vygotsky and Kolb's experiential learning theory was developed by Lave and Wenger (1991) commonly known as Communities of Practice (CoP). With situated learning as the basic principle of this theory, a community is: a group of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis"(Radford et al., 2017; Wenger, McDermott, & Snyder, 2002). One characteristic of those communities, is what Wenger describes as *joint enterprises*. These communities are usually created, developed, maintained by the users, and through this *mutual engagement* they create a unique, to them, social entity (Radford et al., 2017; Rosenbaum & Shachaf, 2010; Wenger et al., 2002). All the participants engage in the activities of a CoP *voluntarily* (Wenger, 1998).

Moreover, the structure of the CoP is another characteristic comprised by a *framework of rules*, *activities* and resources, which evolve through the communication, the *engagement* of the users and their social routines (Baker-Eveleth, Sarker, & Eveleth, 2005). Two important features must be evident in a community of practice: *practice* and *identity*. *Practice* is the first key characteristic of CoP (Wenger, 1998). More specifically in online CoPs, according to Rosenbaum and Shachaf (2010): "the practice of answering questions is the common social practice for the users of these communities." This can be observed through the *mutual engagement* of the users, *joint enterprise* and *shared repertoire*. The *mutual engagement* is observed through the collaboration of the users, regarding problem solving. The guidelines, symbols, and anything of importance to the users is the *shared repertoire* of the community (Wenger, 1998). The second feature, the *identity*, is observable through the participation or non- participation of the users along with the 'modes of belonging'(Wenger, 1998).

According to previous conducted research, CoPs are not exception to any other theoretical learning theory, and thus, do not work for each circumstance of cooperative learning environment. Research showed that they could be hindered by insufficient time for development for the users (Correia, Paulos, & Mesquita, 2010) which in result, lead to insufficient time for trust build-up, and low level of cooperation between users (Radford et al., 2017; Smith, Barty, & Stacey, 2005).

To conclude, the characteristics that a community must share in order to be considered a CoP, according to Wenger (1998) are: the mutual engagement of the users, to create a joint enterprise who shares a set of guidelines and rules who are shaped based on the users' needs, a shared repertoire, as well the users to be the central processing power for the information exchange (learning).

CoP characteristics in CQAs

Previous research who focused to find similarities or if CQAs could be considered as communities of practice has been limited. Indeed, CQAs tend to use the community to create meaningful conversations through which, the information exchange is taking place. There is usually no gatekeeper (e.g. a teacher) that controls the environment except for the users themselves who act as moderators while information is equally distributed from-to the members of the community (Kop, Fournier, & Mak, 2011).

A common theme of interaction, even in the professional environment of Stack Overflow, is that users tend to learn through real-life worked examples (Plass et al., 2010), where sharing previous experiences aids new users in the acquisition of knowledge in an "unstructured" (by formal learning standards) environment.

Previous research has shown that not all members of an online community are willing to share their experiences and contribute in the question answer exchange (Shachaf, 2009). There are members who are working as observers, the "ghost users" of the community who participate through either pure observation, or through anonymous functions such as the upvoting/downvoting system. This percentage of users should still be considered as members of the community according to Wenger (1998).

Additionally, information exchange in online communities has not always been equally distributed with researchers arguing that there can be an imbalance as to who does the most work, who benefits as well as the actions required for this constant distribution to be sustained (Haythornthwaite 2008). For this reason, not every CQA can be considered as a a CoP and it differs based on the platforms ability to provide a sufficient guideline, activities as well as actions through which knowledge will be distributed as equally as possible to all the members. All of them act as "practitioners" since "what they learn from the community affects what they do" (Bates, 2018).

Given that the guidelines in a CoP must change based on the users' needs many CQA sites have included in their joint enterprise extra sites where users can discuss and argue around the shared repertoire and aid to its re-shape according to their needs (e.g. Yahoo! Answers Suggestions Board).

It could be argued that through this participation in CQAs members are able to acquire lifelong learning. Learning is ever-changing and as the experiences increase, for each member of the community, it affects how they react and approach new experiences which consecutively affects the way they learn (Yardley, Teunissen, & Dornan, 2012). Besides the learning shifts, this experience exchange, as well as being an active member of the group, potentially may introduce changes into how the users interact and communicate in the CQA platform, such as Stack Overflow. Changes that reflect in the way the users formulate their posts, their questions, their answers, drastically modifying their "primal" standards of interaction as they climb up the ladder in a reputation system.

6. Method

This section describes the methods used for data collection through the funded university project, as well as my personal selection criteria and analytical methods that were developed for the selected data. Firstly, I introduce the initial data collection conducted by the research project, followed by my personal organization of data based on the answerability of the questions. It is then succeeded by the selection criteria and further organization of the selected questions. In the second stage, I present my personal collection of comment data based on the previously selected questions. Finally, I introduce the mixed content analysis method regarding the questions' formulation, development of the coding framework along with quantitative and qualitative method (latent analysis), concluding with the coding framework developed for the analysis of the interaction between users, in the comment section.

6.1 Data collection and search criteria

In order to investigate how users inquire about information (questions) and how this first step of information exchange takes place between the community (comment section), data was collected through the SO's Application Programming Interface (API).

The data related to the questions, was already gathered for a similar research conducted by an ongoing collaboration project between the *Faculty of Education* and the Faculty of *Applied Information Technology (IT)* within the University of Gothenburg. These questions were initially collected in JSON files which were later transcribed into Excel files. The initial total number of the questions (*N* 205.467) had been previously collected by the funded project. Given the previous programming experience of the researchers who collected the initial data, the Python language was selected. This would be the initial purposive sample of this research. In order to create a stage sample of all the initial data, a specific time frame was selected to further reduce the amount of data for manual analysis. All the python questions posted in 2018, were selected. For each month of the year (2018) a unique Excel file was created.

In total 205.467 questions were collected by the project, through the months January to December 2018. Through my initial effort to familiarize myself with the data I found that the total number of 205.467, was consisted of 130.399 questions which had received an answer and 75.068 that remained unanswered (see Table 1).

Month	Total N questions	Total N answered	Total N unanswered
January	15625	10461	5164
February	15693	10511	5182
March	17611	11662	5949
April	16473	10593	5880
May	17044	11073	5971
June	16576	10722	5854
July	17296	11184	6112
August	17315	11353	5962
September	15778	10305	5473
October	19260	11349	7911
November	19525	11235	8290
December	17271	9951	7320
Median	***17157,5***	***10897,5***	***5955,5***
Total	**205467**	**130399**	**75068**

Table 1: Questions-collected per month overview table

For each question, the following data was previously collected by the researchers:

- Tags
- Owner's (of the question) Reputation
- If the question received an answer (true/false)
- View count
- Accepted answer id
- Answer Count (Total number of answers received)
- Score of the question (According to Upvote/Downvote System)
- Link (to the SO question)
- Last edit date
- Title
- Question body

On the first stage of my personal involvement with the data, a sufficient sample had to be acquired in order to analyse the formulation of questions of SO. Regarding the investigation of the differences and commonalities between a high-ranked an low-ranked and an unanswered question a sample for qualitative mixed content analysis had to be acquired. Overall, stage sampling was used. *"Stage sampling is an extension of cluster sampling. It involves selecting the sample in stages, that is, taking samples from samples"* (Cohen, Manion, & Morrison, 2011).

Three data categories proved to be of most importance for sampling:

- Score of the question (According to Upvote/Downvote System)
- If the question received an answer (true/false)
- Tags

For the second stage of this thesis, as I wanted to investigate the interaction between the users in the comment section, I personally had to collect the comments related to the selected questions. Since the total number of selected questions for analysis was not significantly high (*N* 114), it was possible to manually gather the comments on those questions through the usage of the *Link (to the SO question)*. Hence, no data mining techniques or computer aided software were used for the selection of the comments.

6.2 Data analysis

This mixed content analysis study is going to focus on: 1) the distinct characteristics regarding the formulation of questions being posted by community users in the SO platform, according to the score rating system (upvote/downvote) 2) the discussions developed around the posted problem and its potential solutions provided by the community, in the comment section of the same questions.

6.2.1 Selection criteria

For the first step of the analysis, the questions had to be limited in a reasonable number in order to be thoroughly and manually examined, given that the general aim was to observe the differences between a high-ranked, low-ranked and unanswered questions according to the SO community (see Figure 4). Through the creation of a coding framework, it was possible to distinguish what are the characteristics of each category of questions (based on the *if…answer (true/false) count* and *question score*) and if there were observable differences regarding the askers who formulate each group of questions (according to the *reputation score).*

Additionally, it was possible to examine if the users stick strictly to the SO guidelines for the creation of a good question, or if the community has created its own patterns that can affect either positively or negatively the rating of a question. The differences between users were also examined to observe if users with higher reputation score were able to create more high-ranked questions against new contributing users (with low reputation level) (see Figure 3).

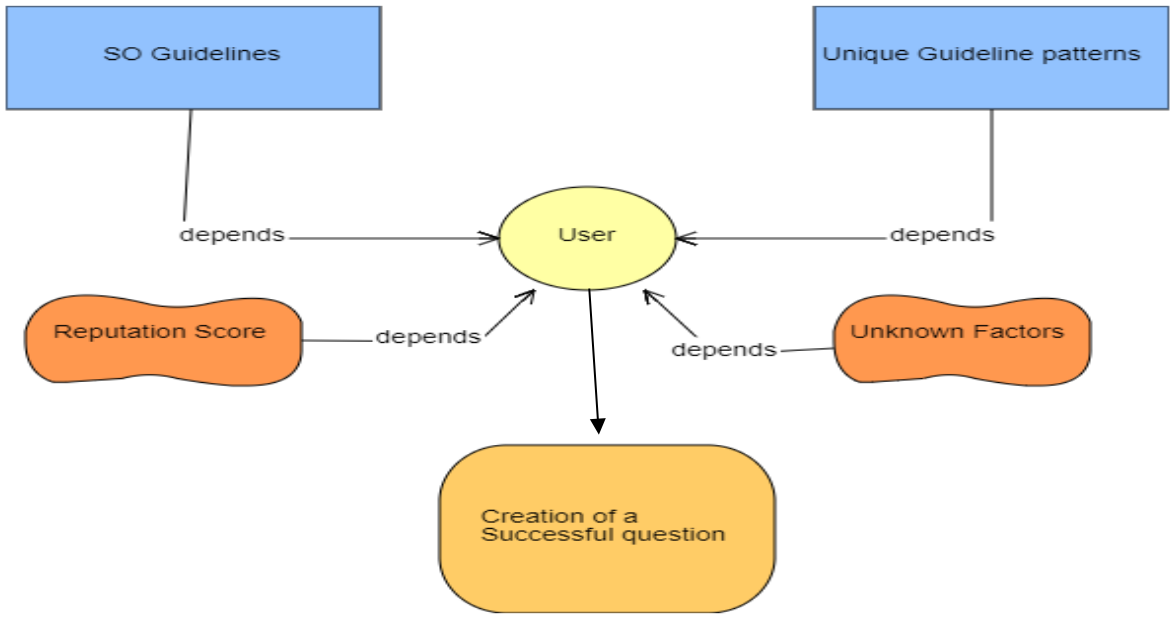

Figure 3: A visualization regarding the possible dependencies for a successful question

For this reason, 3 categories with 2 subgroups were made, which were used to distinguish the different types of question-formulation:

- High Rated Questions (HRQ)
- Low Rated Questions (LRQ)
- Unanswered questions
    - High Rated Unanswered Questions (HRUQ)
    - Low Rated Unanswered Questions

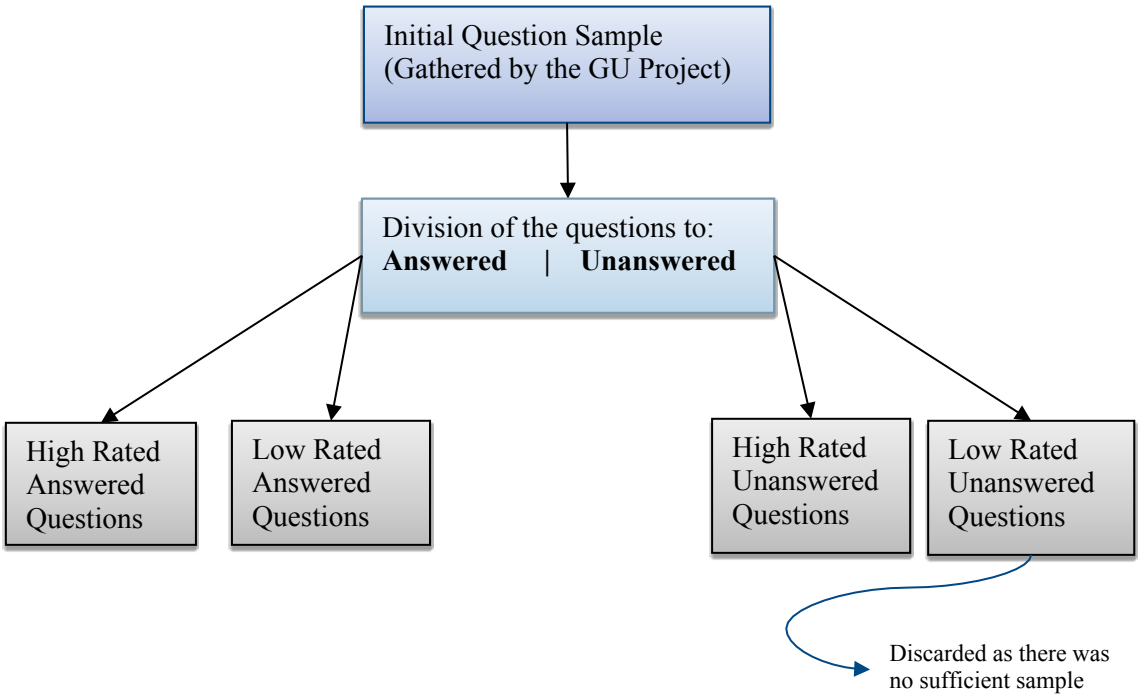

Figure 4: A visualization regarding the first step of the analysis process

Three questions representing each month, for each group and subgroup were selected. In total 36 HRQ and 36 LRQ were selected. This number proved to be sufficient enough to represent the Python community, while descriptive enough to manually analyse and formulate an idea around similarities and differences of each question group. It should be noted that one of the low-rated questions, initially selected for the October month, was deleted until the time of the analysis. In order to maintain an equal number between the highest rated and lowest rated questions, the next in the lowest order was selected.

The initial thought was unanswered questions to be examined as a separate entity of the community; thus, HRUQ and low rated unanswered questions were collected. 42 questions were selected in total from which 36 were HRUQ and 6 Low Rated Unanswered questions. Unfortunately, the number of low rated unanswered questions was the total number available in the platform and from those six, three got deleted in the passage of time. This could possibly show that users tend to delete their unanswered questions when they get downvoted. As of the time of the analysis, only 3 of them were still available online, with a link to the other ones suggesting that the page not found: “This question was removed from SO for reasons of moderation”. The rest 3 that were still available all got merged.

For the second stage of the analysis, in order to investigate the patterns and characteristics of information exchange between users to later compare it to the characteristics of CoPs, all the comments beneath the previously selected questions were collected. The total number of comments collected and analysed was 476. The whole comment section thread was collected which generated the ability to observe the frequency of participation of each user, the interaction among the users as well as the comments made by the initial asker. Time and date of each comment creation were also available.

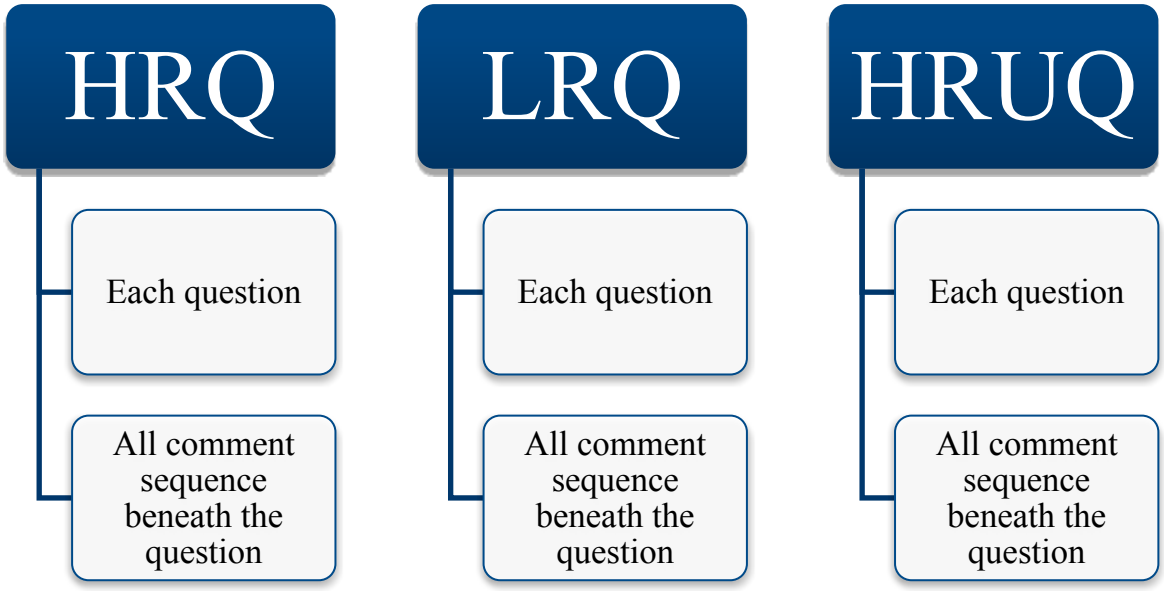

Figure 5: Visualization of the comment sequence analysis

From the 476 comments in total:

- 114 comments were related to HRQ,
- 162 to LRQ,
- 200 to HRUQ

The comments were examined in relation to the questions posted as well as in relation to each other (see Figure 5). This was preferred as the main focus of this analysis was to investigate the patterns regarding the interaction between users and how they aid the asker in the issue through the information exchange.

In both stages of this two-step analysis of data, mixed content analysis methods were used. Content analysis "simply defines the process of summarizing and reporting written data – the main contents of data and their messages while focusing on language and linguistic features, meaning in context "(Cohen et al., 2011)**.** A quantitative approach to observe frequencies was used, as well as a qualitative approach to get a richer understanding of the patterns observed. For the quantitative analysis of data SPSS was used in order to observe the means and medians.

6.2.2 Quantitative data analysis regarding the formulation of questions

Content analysis is defined by its flexibility to adapt to any research design and questions (Harwood & Garry, 2003). The material was analysed step by step, by distinguishing patterns around the formulation of questions, creating categories (of patterns), in order to objectively quantify the frequencies of those phenomena (Krippendorff, 2018),regarding all 3 groups of questions. An inspiration for the creation of the categories was the research of Althoff, Danescu-Niculescu-Mizil, and Jurafsky (2014) which was conducted in the platform of reddit. According to Hsieh and Shannon (2005) directed content analysis, also known as deductive, is usually based on existing theory or prior research and it is utilized to further elaborate on an existing phenomenon. The researchers in this paper examined questions and analyzed what made them successful through the investigation of *factors of success*. These factors included *textual factors* such as: politeness, evidentiality (support requests with evidence), reciprocity (help to be helped attitude), sentiment, length as well as *social factors* such as: status and similarity. The *factors of success* proved to be relatable to the characteristics of CoPs as they represented the mutual engagement of the users (similarity), the way that part of the "shared repertoire" is created in the platform (evidentiality/reciprocity), as well as how the identity of the community users (status) could potentially affect the information acquisition. As Reddit CQA platform differs significantly from the professional formal SO, these categories had to be reshaped in order to fit the SO platform and produce findings relative to this study's research questions. An emphasis as thus was put in the *evidentiality, sentiment and status.*

At a first stage all selected questions were observed to explore the usage frequency of the aforementioned factors of success (patterns), which were related to the SO guidelines regarding the formulation of questions. After many iterations, categories were well-defined and usage was observed, while a latent content analysis was conducted to infer meaning, interpret the content which was previously categorized (Hsieh & Shannon, 2005).

On a first part general categories were created. These categories were: **Explanatory text, Code Snippet, Grammar, Comments/ Edits**. This initial framework followed the most observable patterns (4 in number) which seemed to be closely related to the SO directions or what Wenger (1998) describes as "guidelines", and one specific category regarding the **comments and edits** (community acting as a joint enterprise). The latter was chosen to quantitatively investigate the discrepancies between the different groups, in relation to the number of comments and edits they collected. It would also give an insight as to how users are able to aid the asker through the re-formulation of the question in order for this "shared repertoire" to fit the community guidelines and be available to other users in the future. As a result, categories and the names of those came through the analysis of the data (Hsieh & Shannon, 2005). Accordingly, whole sentences were classified into much smaller content categories (Weber, 1990). Through the first iteration it was evident that more distinct categories should be created, which would focus on the title creation, **Tags, Feelings** (instead of grammar), as well as **external resources** (which previously were part of the explanatory text). Regarding the Evidentiality, one category was created around the **structure of the question body**, which consisted of 5 subgroups: **Explanatory text, Code Snippet, Mention what has been tried so far, clear distribution**, as well as a **subgroup that entwined all of the aforementioned features** to distinguish questions that included all of the SO guideline directions. As thus, it was possible to combine interacting categories to create a whole, general idea (Chinn & Kramer, 1983; Hsieh & Shannon, 2005) around the key main features of well-formulated, high-ranked questions and what distinguish them from the LRQs or HRUQs.

This step-by-step formulation of deductive categories, made possible to manually observe the frequencies under which, the users urge to follow particular patterns throughout the formulation of the question. After the two revisions, the level of abstraction was reduced, creating distinct, subgroups for those 6 initial codes, which were developed and examined as separate entities (Hsieh & Shannon, 2005). The final groups and subgroups were regarding:

1) Format of the initial question (heading)

Code	**Definition**	**Text language example**
a) In the format of a question	If a heading included a question mark and/or had a syntax resembling a question	"How do I operate on a DataFrame with a Series for every column (?)"
b) in the format of a title	A heading resembling a normal, neutral sentence.	"Python Pandas User Warning: Sorting because non-concatenation axis is not aligned"

2) *Structure of the question body*

Code	**Definition**
Following overall the structure suggesting by the SO guideline	Inclusion of both "explanatory text", "code snippet"", "what they have done so far"
Inclusion of explanatory text	Summarize their problem with the inclusion of details about their goal, describe expected and actual result with the inclusion of any error messages
Inclusion of a code snippet	Provide some code snippet, so that others can observe the issue at first hand and replicate the code, if needed.
Inclusion of what they have done so far	Describe what they have tried so far
Inclusion of **bold,** *italics and having generally a clear distribution*	Users that manipulate the font, the style, usage of bullets, to create clear distributed questions

3) *Expression of feelings*

Questions where users do not use the neutral tone which is considered the normal, suggestive way of SO. This could be indicated with the usage of words or phrases that express either gratitude (*"thanks in advance"*), distress (*"…its driving me mad"*), sympathy ("*I do not come from a coding background so bear with me…*".

4) *Inclusion of tags*

In three distinct categories (see Figure 6) it was calculated the Total (N) mean number of tags used in a question group (e.g. HRQ), the 2nd most used tag after "Python" in this particular question group, as well as a distinct subgroup which included questions with only 1 tag.

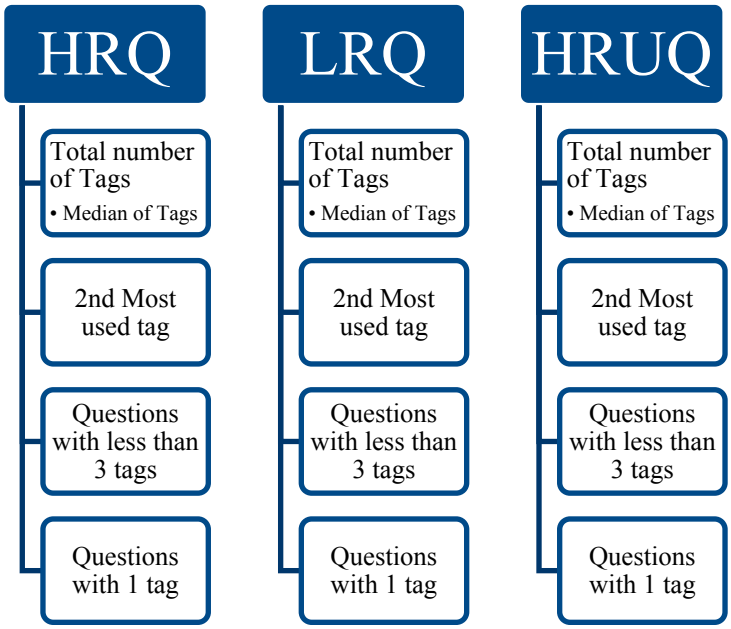

Figure 6: Visualization of tag analysis for each question group

5) Linking to *external resources*

The SO guidelines include in the same suggestion that the users should include an explanatory text AND the inclusion, if possible, of any information found on the SO site (previous posted questions) or any other external link and why it didn't meet their needs. It was observed that even though users tend to include an explanatory text, the utilization of external resources was unrelated to the explanatory text, with many users never providing external resources. Thus, a distinct category was utilized to observe, which, why and how those users utilize the external resources. During the analysis each question group was examined as to:

- how many external resources was able to gather (N)
- the median of the external resources
- Links to previous posts, comments, answers in SO
- Links to python dictionary
- Links to Additional Documentation
- Links to tools and websites related to the code provided

6) *Comments and edits*

In the first sub-category it was observed a) the total number of Comments received in each particular group (e.g. HRQ), b) questions that received zero (0) comments. In the second sub-category it was observed a) total (N) of questions that got the final edit by the asker, b) total (N) of questions that got the final edit by another user, c) total (N) of questions that did not get an edit either by the asker or another user.A deeper analysis took place to create a more detailed image of how the platform and the community users interact and aid the asker through the edits. The analysis was done manually by reading and creating another coding framework around the number of edits, and what was the main focus of the edit for each group of questions. On the first-step, this led to the creation of the table below (See Table 2).

	HRQ	LRQ	HRUQ
Edits regarding the title			
Edits regarding the restructure of the main body			
Edits regarding additional text			
Edits regarding typos and grammar			
Edits regarding tags			
Edits regarding bounties			
Edits regarding code			
Edit to fit SO guidelines			
Edit to remove text			
Total number of unique edits			

Table 2: A representation of the 'Edit' analysis regarding all the question groups

6.2.3 Latent content analysis in regard to the question formulation

According to Elo and Kyngäs (2008) qualitative content analysis wants to preserve the advantages of quantitative content analysis for a more qualitative text interpretation. As such, observational findings that affect the formulation of the questions but cannot be thoroughly defined through a simple quantitative process, were analyzed. As such, the user's median reputation for each category of questions, an interpretation of the results, as well as "extreme cases" were discussed in the qualitative section beneath each separate category classified as *"Observations"*. At the same time a table was created to investigate the correlation between the mean score of the overall questions and the reputation of the users. Distinct categories were created to investigate the number of low reputation users, which ranged from 1-100, 101- 1000, 1001-3000 reputation score. At the same time the highest reputable users found in each category were mentioned, to observe if extreme cases of high reputation are found in each question group. All the above were analyzed based on the format of the initial question of the users. A form of *concept analysis,* which is frequently used in nursing studies, was used to develop understanding of the meaning of communication, intentions, consequences as well as the context (Downe-Wamboldt, 1992).

6.2.4 Mixed content analysis of discussions in the comment section

For the second step of the analysis a separate deductive focused, content analysis was used. As thus, for the creation of the initial table the analysis was based on the coding framework as presented by Haythornthwaite et al. (2018), researchers who investigated subgroups of Reddit and conducted a content analysis. Their table was based on Ferguson, Wei, He, and Buckingham Shum (2013) evaluation around exploratory dialogue cue phrases.

The prementioned table proved fruitful for the prediction of variables of interest, the relationship among those variables as well as the help to determine the initial coding scheme (Hsieh & Shannon, 2005). According to multiple researchers the way of formulation of an answer, or comment in this case, as well as the transaction level can differ from site to site levels (Harper et al., 2009; Rosenbaum & Shachaf, 2010). At the same time the coding framework should reflect the research questions of this study. The focus of this analysis, was related to the usage of the *community* to produce *learning* through information exchange and if its procedure could be related with the central role of the *community*, as described by Wenger (1998). As thus, the categories reflected the importance of the community's *guidelines*, how the active users shape the *shared repertoire* through *mutual engagement*. In order to be able to distinguish those CoP characteristics, emphasis was put on the discussion around the formulation of question and discussion around the potential solution *(shared repertoire),* common goals and argumentation against potential solutions (*mutual engagement),* as well as reminder of the SO guidelines and the communities framework regarding the formulation of questions (*joint enterprise and framework of rules)*.

Few patterns developed by the team of Haythornthwaite et al. (2018), proved to be not significantly present in the SO platform, as the level of "socializing" is dramatically reduced in comparison to an informal platform as Reddit. Observable patterns of SO that were not able to be categorized based on the previous categories, were identified with the representation of it through new categories or subcategories (Hsieh & Shannon, 2005). Thus, new definitions and examples were given to aid the user's understanding behind the categorization and the frequency of patterns observed (see Table 3). The examples used to describe the patterns were obtained equally as fragments from all 3 question groups. They were evident in all groups hence, they were considered patterns, with only exceptions being *common goals* and *Discussing the question in a more negative manner* which were not evident on all three groups.

Code	Definition	Text language examples
1. Argumentation	When the asker or another user, argues against a proposal made in a previous comment by another user/asker.	"your printm() is wonderful but…" "This is not a dupe as my issue focuses mostly on..." "Is this really a duplicate?" "This, does not answer my question"
2. Argumentation against formulation	When other users criticize the code, main body formulation as posted by the original asker	"You should switch to Python 3.6 while you are still learning", "why have you not even attempted it?"
3. Inquiry for further explanation in a neutral manner	When another user, is neutrally asking for further clarification to the original question, or further evaluation to the methods, provided in the comments	It would be nice if you could provide further code, thank you!', "Have you tried setting the TensorFlow….report the results?"
4. Inquiry for further explanation in a positive manner	When another user, is positively asking for further clarification to the original question, or further evaluation to the methods, provided in the comments	Please provide us with the…. Thanks", "What OS are you using? :)"
5 Discussing the question in a neutral manner	Discussion around the solution to the issue at hand with neutral linguistic expressions, inside the SO guide norms	This is just a guess but...', 'Have you tried...', 'It would be possible to...'
6. Discussing the question in a more negative manner	Discussion around the solution to the issue at hand with negative linguistic expressions, outside the SO guide norms	"Have you even tried to read the description?" "How is someone supposed to help with no feedback on..." "Have you even tried to solve it on your own?"
7. Discussing the question in a friendly manner	Discussion around the solution to the issue at hand with friendly linguistic expressions, outside the SO guide norms jokes, funny expressions, "puns" are also included (sometimes it includes emojis that embrace the asker's frustration)	"Thank you for the heads up." , "Or the same if the computer crashes before :D", "finally will not execute if the power cord is ripped from the wall."
8. Common goals	Comments from users that are facing or have faced similar issues in the past, with possible solutions	"Any solution to this? I am facing the same issue" "Yep, same problem"
9a. Additional information	Comments that provide links to external resources, in an effort to provide additional information	You could take a look at this…' "There is a (github) that might solve your problem"
9b. Additional information (Site specialization)	Comments with links to internal resources, either in SO or other Stack platforms	"This could be a dupe of your question" "Could this question solve your issue?"
10. Reminder of SO principles	Comments that have as a common theme to promote compliance to the SO principles this could include comments that only mention "this is a dupe/lacks clarity/ not focused" without further explanation as to why.	"This is a duplicate [link]", "The description lacks clarity" "This is not a coding service..." "Please refer to the S.O guide… before posting"

Table 3: A coding framework representation regarding the analysis of the comment section.

Through the iterative process of observation and analysis, it was observed that certain comments would fit in multiple categories as they would yield features of e.g. argumentation and additional information. Those comments were distributed to all the corresponding categories that shared features of.

6.3 Ethical considerations

The dataset used in this study was gathered online through the SO's API. Even though it is an open site and SO does provide the researcher with a guiding tool as to how to mine the data, some ethical considerations should be addressed. The funded project (SOCDEX) was approved by the Swedish Ethical Review Authority (in Swedish: Etikprövningsmyndigheten). According to the proposal made by the researchers: "the project was conducted in accordance with the Swedish law on ethical vetting for research involving human subjects (SFS 2003:460)" while at the same time "all personal data collected was handled in accordance with the incoming *General Data Protection Regulation* (EU 2016/679)".

During the analysis, an effort was made to protect the anonymity of users that were involved in the creation of the question thread as well as edits, comments, flagging. Online personal data of the participants was observed and used in the form of their reputation scores through the creation of means and medians in this study attaining simplistic results (Weber, 1990). Additionally, extra focus was put to two of the highest reputable users at that time observed in the data sample, in order to analyze some of their unique patterns and how high contributed users of the platform formulate their questions. Even though there was an effort for all the data to remain anonymized throughout the study, those users could potentially identify themselves through the mention of their score as well as "writing style". Given that the biggest part of the analysis had qualitative characteristics, individual cases and instances were examined. An effort was made so that the principle *primum non nocere* -do no harm to participants- would be addressed (Cohen et al., 2011). Finally, for the creation of the coding frameworks, text examples were used so that the reader will comprehend the different categories as well as the data that corresponds to each one respectively (Elo & Kyngäs, 2008; Patton, 1990). Even though an effort was put to keep the examples short and utilize the most frequently used text citations ("quotes"), some users might be able to identify themselves. For the same reason, direct longer quotes were not used during the latent content analysis of the findings, even though it would potentially boost the credibility of the study.

Another potential issue could be in the form of bias. I was personally involved in the funded project by the University of Gothenburg during my internship. Moreover, I have been a training programmer, specifically in Python language and I was a passive user of Stack Overflow, few months before I started this thesis. This could potentially have affected the selection of patterns and hypotheses that I have laid out throughout the thesis. On the other hand, it could have potentially aided in the detailed development of the empirical context as I was personally involved, even as a passive user of the platform.

Finally, an emphasis should be put that the data was mined from the website. Even though it is a common practice in the analysis of CQAs it should be noted that this data was not created with a researcher in mind (Cohen et al., 2011). Stack users are potentially not aware that their questions and comments could be utilized as a sample material for research and given that Stack provides the opportunity for personal information such name, reputation score and direct links to the question at hand, this could even further create additional ethical considerations. Knowing that there is actually an ongoing discussion around ethics regarding social platform mining data, an effort was made to keep the sample as anonymized as possible.

7. Results

In this section I provided the results that I was able to gather through the two-step analysis process. In the first step, regarding the formulation of questions, I separated the results based on the categories/patterns that were analyzed, independent from each other. I have merged the quantitative results of each category, for each of the three distinct question groups, in order to make meaningful comparisons and observe the differences between the groups. Sections marked as *Observations* indicate the latent part of the content analysis, where I present qualitative results which became apparent through the observation process and which were hard to portray in a quantitative approach. On the second stage of the analysis regarding the interaction in the comment section, I follow a distinct approach. The results, either qualitative or quantitative, are linked but each code category is examined separately by investigating the commonalities and discrepancies between the question groups.

Through the mixed content analysis and the application of the coding schema developed, the formulation of questions showed some discrepancies between the 3 different distinct groups. The most frequent observable difference was regarding the connection between the question's score and user's reputation. Even though there were common patterns existing in all groups in relation to the question formulation, a significant difference was observed regarding the participation of other users through edits and comments. This created the interest for the second step of analysis for this study (comment section).

7.1.1 Format of the initial questions (heading)

It seems that the format of HRQ, composed as a question or as a heading, did not show difference in frequencies in HRQ (see Figure 7). Similarly, the LRQ showed an equal distribution among the questions formatted as questions or headings (see Figure 7). A difference in style was seen in the HRUQ where askers seemed to have preferred a formulation through a title. In order to investigate further this phenomenon, a deeper qualitative analysis was done to observe the mean score of the questions and median of the reputation score of the users (see Table 4).

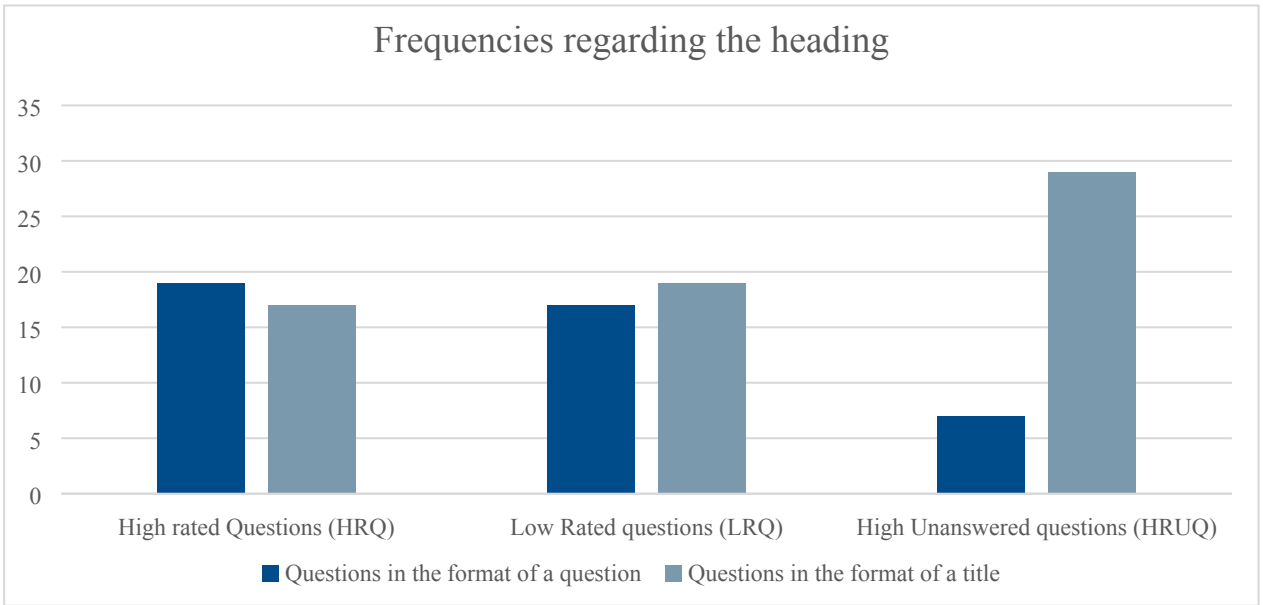

Figure 7: Initial question formulation: Distinction between Title and question-like format

7.1.2 Asker's reputation and question score

HRQ users that formulated their questions as a question had a median reputation score higher than those formulated in the style of a title (see Table 4). The difference between the mean question score was not that significant especially for the LRQ and HRUQ. Indeed, the HRQ formatted in a heading style led to a mean score of 135,6 as opposed to 86,6 but the reason behind this was found to be related to the number of users attracted by the matter at hand (*a more niche question will attract less users even if it is formatted in a heading style*). Through the calculation of reputation of users, it was found that 5 out of the 36 HRQ were created by the highest rated users (197.000 and 203.000 respectively). Half of those questions were created in the format of a title while the other 2 in the format of a question.

The mean questions' scores of HRUQ group were not different but an increase in median user reputation was noticed with the utilization of the format of a question. Unfortunately, the number of HRUQ askers that preferred the question format was significantly less that the other, thus the increase in overall question score cannot be said if it is related to the initial heading choice. Surprisingly both LRQ question formats had significantly lower score than all the other groups. Even though the mean questions' score was quite similar, the median reputation score showed an even lower number for the users that utilized the title format.

This led to the creation of further categories to observe how the users of all groups are distributed between the 3 level of low reputation. HRQ and HRUQ had almost zero users that belonged between 1-100 reputation score while the majority of the users in LRQ were part of that reputation group. In regard to HRUQ, it was evident that almost half of the askers, that utilized the format of a title, belonged between 101- 1000 reputation range. This table also showed that both groups of HRQ had significantly less users beneath 3000 range in comparison to LRQ and HRUQ.

	HRQ in format of a question	**LRQ in format of a question**	**HRUQ in a format of a question**	**HRQ in a format of a title**	**LRQ in a format of a title**	**HRUQ in a format of a title**
Mean Question Score	86,6	-10,5	13,4	135,6	-10,2	13,2
Median Reputation Score	4473	11	900	2263	3	576
Users in range of 1- 100 Reputation	0	12	0	0	17	2
Users in range of 101-1000 Reputation	5	3	3	3	1	14
Users in range of 1001- 3000 Reputation	3	2	2	4	1	4
Highest users present (Reputation Score)	1= 203.000	1 = 20.000	1= 75.000	1=197.000	1=1400	1= 61.500

Table 4: Further analysis on the initial question formulation among HRQ and LRQ

7.1.3 Observations regarding the Format of the initial question (title)

The initial drawbacks of the LRQ showed up from the title formulation. Most of the questions did not have a question-like format or title similar to the others. They were mostly formatted in *dictation*. Additionally, many of those users lacked in clarity even in the formulation of the initial question. Naturally, questions with a more well-formulated initial question, that followed the community guidelines, got less downvotes in comparison to the rest. These downvotes might came from the whole structure of the question, missing code, explanatory text, etc. categories which will be analysed later on.

Only 6 of the 36 of HRUQ were formed in a question format (see Figure 7). Most of them were utilized a title having the main theme of the overall question as a single sentence. Sometimes users tended to use the initial framework that they are working on before setting the question (e.g. Facebook Graph API- followed by the title). One pattern observed was that headings were filled with information in comparison to HRQ titles. Fewer times the title was big enough to require a second row. As SO insists, the creation of the title is what attracts users to read the question body. HRUQ askers seemed to failed to capture in one sentence the issue that they are facing, providing too much initial information, which could potentially create confusion to other users.

7.1.4 Structure of the question body

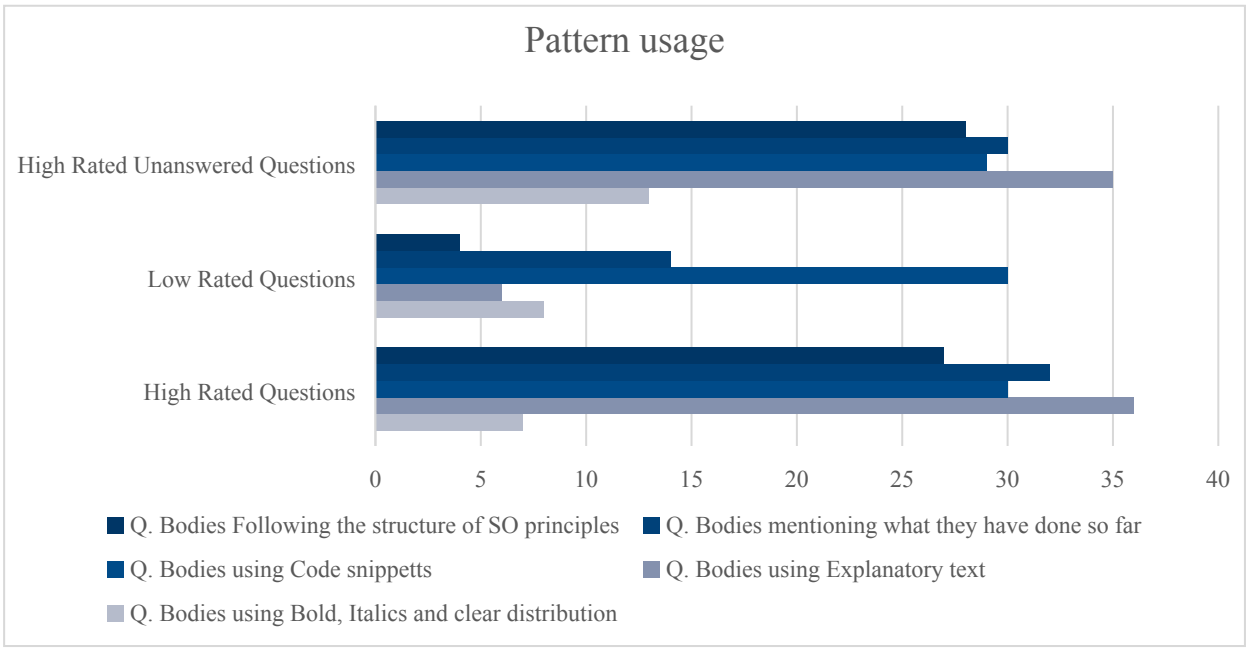

Figure 8: Common patterns around Question bodies

Utilization of Bold, Italic and clear distribution: The use of "tools" regarding the structure and clear distribution of a question was the least used pattern regarding the formulation of a question body. The biggest percentage of utilization was observed by the HRUQ. HRQ and LRQ showed equal numbers of utilization. It should be noted, that even when LRQ followed the use of bold, italics, it was done less methodically and with less clear distribution comparatively to HRQ.

Utilization of an Explanatory text: Explanatory text is considered one of the most important foundations for a successful question in the SO community as the asker is usually required to explain what they are trying to achieve and their knowledge around the issue. Unsurprisingly all the HRQ askers (36 out of 36) included explanatory text to their questions. Accordingly, HRUQ (29 out of 36) had explained what they are trying to achieve through descriptive text. Another difference was observed in the LRQ where the majority of the users (30/36) did not use any form of explanatory text. This is considered a big misstep from the community, as it will be thoroughly analyzed on the second stage of this study.

Utilization of mentioning what they have done so far: At the same time 32 of the total number of HRQ askers, provided *what they have done so far* in the question section, while 4 that did not it presumably was related to their initial question being more conversation focused. The HRUQ followed closely to that regard with 30/36. LRQ failed to do so in the majority of the questions, with 22 having zero text related to the previous efforts.

Utilization of code snippet(s): The code snippet seems quite powerful for collecting meaningful input from the community. All the HRQ code snippets were again according to the regulations of the SO. The HRUQ group followed in the same steps but there were exceptions. Seven out of the total 36 askers, used code screenshots or external links for their code. That is considered to be against to the SO and community guidelines. The same pattern was observed in LRQ in a bigger number of questions. The community does not agree with the provision of screenshots as it might lead to some users not having access to them. Even though it might not be correlated with the inability of the questions to receive an answer, or with their low score (in the case of LRQ) it was one of the contrasts in comparison to HRQ.

Following strictly the structure guidelines provided by SO: The majority of HRQ and HRUQ users followed the structure of a question as suggested by SO community guidelines (see Figure 8). Thus, setting up a title, and creating a body which is made up from a brief explanation of the issue at hand, a code snippet, a/some sub-question(s), what have been already tried, what is the error/ or what does the user try to achieve. Even though 9 out of 36 HRQ were not following all of the rules, some of them had a conversational style (thus no code snippet was provided) and 3 were following a simpler, though clear, distribution in comparison to what SO suggests.

7.1.5 Observations regarding the structure of the question

Almost all the high rated questions (answered and unanswered) included another question into their text. Sometimes this question was closely related, or more like a rephrase of the initial question, while other times the asker used the initial question by generalizing the issue at hand and through the question body, with 2 or more consecutive questions, tended to unlimit the issue, especially if the issue tended to require a multilevel solution. Another common pattern found in some of the HRQ was that

the deeper analysis of the question happened after the first transcription of the code snippet. After that, askers provided additional code snippets to further describe as to what the initial code gave as an error while other users tended to describe their issue with words. Some users urged to use the "code snippet" function to emphasize certain words. This could be observed when users tended to greying out specific words that are code-related (e.g. Series, DataFrame). This sometimes created a more concise and short question which might seemed more attractive to other users as it leaves many chunks of code out of the equation. Even HRQ questions which did not include any code snippets at all, they tended to use the grey text to emphasize certain code specific functions or results.

A more widely adapted usage of *italic* was related to solutions or tools that the asker has considered before but do not really suit their needs while **bold** might be used to highlight important software features that are specific to their issue. Fewer times, the users were likely to use **bold** to present their, what as previously mentioned, research questions, by making the main question split into smaller chunks. In comparison many LRQ posts missed the structure that was evident in the high rated ones. Many users favoured to explain what they wanted to be achieved but not what had already been tried.

Another reason that few LRQ failed, was of not being able to adapt to the community guidelines (see Figure 8). It was quite evident that the users with low-voted questions were new programmers or learners of the Python language. During the formulation of the post, few times they opted to explain what they did so far. Other times requested from other users to solve their issue in one particular way, even if they previously recognised a plethora of solutions to their issue. Given the simplicity of the tasks presented, the users did not face the issue of *high order* in Python.

Even though some of the categories were able to show significant differences between the formulation of the HRQ and LRQ questions, one proved to be misleading. More specifically, the use of code snippets. Indeed 30 out of 36 HRQ used code snippets. The 6 that did not, were conversational questions, where the issue did not lie in a particular coding effort but something regarding specific software or hardware. The same number showed up in the LRQ with 30 out of 36 using code snippets. The contradiction here was that, 4 LRQs that did not use this particular function were created by extremely low (1-10 level) reputation users who never put any effort at providing some code or trying to solve their issue.

7.1.6 Expression of feelings through text

As previously mentioned, SO is a site made from professionals, for professionals, thus it promotes a neutral tone for the questions/comments/answers of the community. Indeed, very few HRQ askers (4/36) favored to express feelings in the form of distress or gratitude. Even when that happened, it was in a very mild tense, with adjectives that tended to describe their situation ("*I feel hazy about…, perplexed, etc."*). There were also fewer examples where askers leaned to show gratitude by appraising the previous posted questions around the same issue, or answerers' suggestions ("*interesting!, that is certainly unique, I am surprised but it could work, that is exactly what I was looking for, thanks"*). As for the HRUQ, 4/36 aimed to use some friendlier expressions (*"any help would be greatly appreciated, thanks*") but with the rest of the text following the manners of the community. It cannot be determined that the users were judged hardly for that, as the question scores on those questions, were quite positive.

Comparatively, 15 out of 36 LRQ askers, did not only show more gratitude in their text, but they urged to thank in advance in the question body. This will be analyzed later on, through the edits, as a large majority of edits was taking place in order to remove those kinds of expressions by other users of the community.

7.1.7 Observations regarding the Expression of feelings through text

Users of the community tended to follow strictly the rules of the community and reversed from informal and friendly attitude among other users, according to the SO suggestions. Users who had created LRQ leaned to express more feelings even during the creation of the initial question. In more detail, some of them were greeting other users as someone would do in a social media group and were apt to express frustration among others ("Hello, I have an issue…"). They further tended to express feelings of understanding as ("it might be stupid but", "I know this is simple", "I know my goal is simple") among others, expressions that were not apparent in the high-rated questions.

7.1.8 Tags

Even though the majority of the HRQ users were trying to engage with as many tags as they were allowed by the system (5 tags), 8 questions, with 4 of them being made by the highest rated users (195K + rating), included only 2 tags. The second most popular tag (after Python) that many other users utilized was to specify the version of their language, with the biggest margin using "Python 3.x" while others aimed at specifically a particular version ("3.6", "2.7", etc.). According to the analysis of the tags collected from LRQ and HRUQ neither of the groups used a specific language tag.

7.1.9 Observations regarding the Tags

There was not really a consensus around tags. Even though HRQ leaned to use at least 3 to 5 tags, there are few HRQ with only 2 tags. Of course, the number of tags means absolutely nothing if the question at hand is not related to the tags or if the tags are not very specific. Most users tended to use important attributes of the question or phrases (gpu, Nvidia, etc.) into the tags, along with more specific and specialized tags such *memory-management*. In comparison, disparities were observed in the tag usage as well as choice of the tags.

The python language has evolved considerably since the 2.0 Version, which resulted in many askers utilizing the python-3.x tag as well. One distinction, between HRQ and for both HRUQ and LRQ was that the language specific versions were missing from tags. This could potentially draw less dedicated users as the answers that they might provide could possibly be not functional in the asker's version. Though it should be noted that given the statistics gathered by SO the Python tag has been utilized 70% more than any other specification of itself.

Regarding the LRQ some users, gravitated towards tagging multiple languages in the tag section which along with the lack of clarity, created even more frustration. Tags were not only sparingly used in LRQ but misused as well. Many times, the asker used wrong python tags (python 3.0 instead of simple python or python 2.0) which led to a wrong part of the community trying to provide an answer. The tags were also less specific.

7.1.10 Comments and Edits

The HRQ were able to collect 114 comments in total with some questions (43%) having 7+ comments. More considerable presence of comments was found in the LRQ and HRUQ where the users would discuss around the main issue at hand. Only 7/36 HRQ and 4/29 HRUQ showed questions without a comment.

Even though the number of total edits in the HRQ (*N 65)* might closely resemble the findings in the LRQ (*N* 57), the reason for the edit varied significantly (see Table 5). One of the easily observable differences was that 10/36 LRQ never got edited, thus no high(er) reputation user put any effort in

redirecting the question to become more attractive to the rest of the community. A similar considerable number was observed in HRUQ (9/36). Moreover, a percentage of HRUQ only got edited by the initial asker with not editable feedback from the rest of the community.

	HRQ	LRQ	HRUQ
Edits regarding the title	10	9	2
Edits regarding the restructure of the main body	9	13	3
Edits regarding additional text	10	8	13
Edits regarding typos and grammar	9	4	8
Edits regarding tags	12	4	5
Edits regarding bounties	5	0	12
Edits regarding code	3	14	10
Edit to fit S.O guidelines	3	3	4
Edit to remove text	4	2	3
Total number of unique edits	65	57	58

Table 5: Deeper analysis of editorial work among the community

7.1.11 Observations regarding Edits

It is important to note that almost all of the HRQ get edited at some point. This point in time might differ greatly with some of the questions getting edited in the same week while others get edits after a whole year, sometimes more. The edits were either accomplished by users of the community with high overall rating, while fewer were the cases where the asker will edit their question.

The way that the editing was happening was another contradiction between the questions. In high rated ones (answered and unanswered), the edits were more precise, fixing minor issues, or adding extra text for further clarification, along with updates, notes, external resources. The titles were revised to be more explicit, while the restructure was mostly focused on using bold, italics and rephrasing parts of the original question.

The edits in the LRQ were quite substantial. This showed from the fact that even though 10 questions got 0 edits, the total number of revisions reached 57. This number almost equalled the 65 of the high rated ones where 33 questions got revised. LRQ askers that did not receive any edits at all had the lowest possible score, with 6 of them being new contributing users (Reputation level 1) of the platform. Askers that have posted few series of code with no, or a single line text, were thoroughly edited by the users, adding additional explanatory text, rephrasing the code and the question body question. Since most of the askers were new users, regarding their contribution score, there have been lots of edits restructuring main bodies, removing unnecessary friendly expressions and feelings, to fit to the SO guidelines. Even though tags were “suffering”, no one pays attention to them during editing. Only two times users removed irrelevant tags as html, javascript.

Regarding the HRUQ, 26 out of 36 got edited. Distinctively, a big percentage (12/35) got only edited by the asker. Consecutively, it could be considered that those questions did not attract the same number of high reputation users, regarding edits in comparison to HRQ. The askers though, showed determination and dedication in order to get an answer to their question. All the questions that were edited only by the askers had significantly more edits. Another difference that could be observed in relation to the HRQ, would be the overall increase of the bounties (see Table 5). 12 of the total number of edits in HRUQ were related to the creation of bounties. Given that the users were mostly in the lower scale of the reputation system, they were unable to reach the bounty levels observed in the

HRQ (where users could bet 250 and 500 reputation accordingly). All the bounties in these particular questions were of 50 reputation but nonetheless, none of those questions received an answer. It should be noted though, that the impact that those bounties would have, to the asker's overall reputation score, would be more significant than of the High reputation askers of HRQ.

7.1.12 External resources

Externals resources usage varied from post to post and from user to user. Thus, they were coded in 4 distinct categories around the usage of those links.

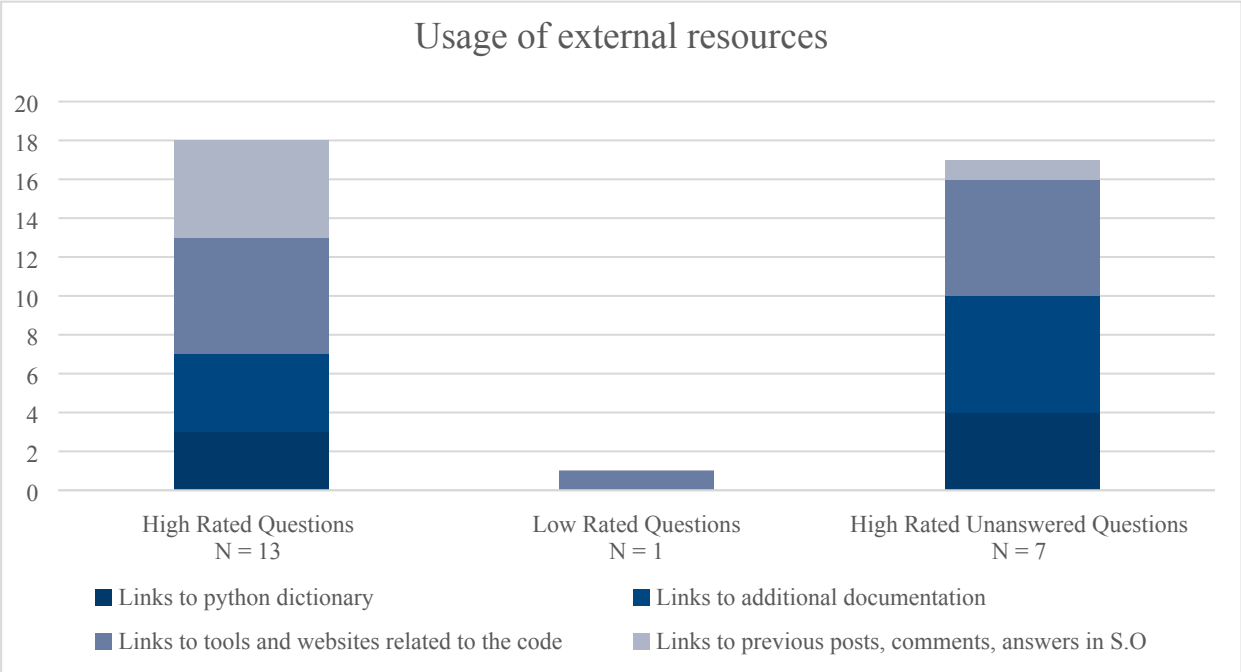

Figure 9: External resources used overview

HRQ showed the biggest usage of external resources with the majority being external links, related to the overall understanding of the presented issue (see Figure 9). The most common links referred to GitHub files and other similar webpages and were closely related to their code. The rest of the external links could be seen as "sources" were the users would provide links to either python "dictionaries" or specific software and hardware manuals. Usually were utilized as additional documentation, an explanation as to why they have formatted their code in that particular way and why they were looking for one specific method for solution. Similarly, 13 out of 35 of HRUQ use resources of different attributes.

The number of links was dramatically reduced in LRQ where only one user in a single post used an external link to inform the rest of the community around the overall matter at hand.

7.1.13 Observations regarding the External Resources

As observed in the figure 9, the usage of external resources in the HRQ was substantially more than any other group. Some of those questions included "items" that the askers wanted to use in their code or small examples of what they were trying to achieve (E.g. *a user wants to learn how to resize a picture to a smaller scale and gives to 2 varying pictures to visually present what they are trying to achieve)*. Sometimes the links they provided in HRQ was regarding courses or external resources that they tried to use to solve their issue. Fewer times can be observed the creation of "word-plays" by metaphorically using items to describe the situation at hand (e.g. a user used the phrase "this is the Swiss Army knife of…" and replaced the word of *swiss army knife* with a link at wiki). This could be seen more often on the comment section where, as mentioned before, users tended to have a more informal tone.

Another use of external resources is when the user suggested other threads or websites that they had read, before taking the decision to write the current post. This way they were able to save users from the trouble to tag and suggest previous questions that might solve the asker's issue.

Even though the HRUQ group was with quite lower utilization of external resources, the format of those link and the distribution of those, was similar to the HRQ.

Finally, only 1 external resource was used in the LRQ. This did not come as a surprise as the majority of the questions were made by new, low-skilled programmers. Thus, the exclusion of external resources did not reduce the validity of the questions at hand, as the usage of those would be considered redundant.

7.2 Comment content Analysis

The total number of HRQ comments analyzed was 114. The final total number of coding patterns observed was 118. Subsequently, the consistency of the comments was quite high, as the majority of them focused on only one attribute of discussion at a time. In comparison, the comment section of LRQ and HRUQ was significantly more populated in comparison to HRQ, showing significantly more contribution by the community, as they managed to collect 162 and 200 comments respectively. Except for the differentiations in frequencies, the consistency of the comments differed as well, as a bigger number of comments belonged into multiple categories (see Table 6) which is probably related with the comments being more extensive (bigger in length).

The first image someone would get by analyzing the comment section is that, the utilization of it, resembles the form of a troubleshoot chatroom. Users ask for additional information, discuss around the matter at hand and provide possible solutions and check-ups to the asker until they find the solution to their problem. This came in many formats. The largest percentage of the comments focused on creating a discussion around the issue of the asker in a neutral manner as suggested by the guidelines in SO (see Table 6).

At the same time a pattern emerged that was considered unique to the rest of the coding framework and thus, was analyzed separately. The pattern was actually in a form of a comment that was used in the exact same format (or with some slight iterations). This passive-aggressive style of a comment was targeting users that did not try to solve the issue posted on their own.

It proved to be one of the most observed comments repeated by users of the community, sometimes more than one time in the same question. The main focus of the comment was: *"What have you tried so far? Stack Overflow is not a code writing service, we're here to help but you'll need to show us what you've tried already"*. Especially the first part of the sentence was repeatedly observed, mostly evident in LRQ posts. After some investigation, it was found that this phrase has been discussed a lot in the *meta stack*, as well as popular blogs and other coding websites from SO users, and the view of the community is divided. The main reason for its "creation" is that there is not an exact *flag* to report those questions as non-stack compliant and many users would want to have a flag with that exact phrase. Even though the majority of the users in those threads seem to be quite negative around askers that pose that kind of questions, (arguing that many of them offer PayPal money to anyone that solves their issue), other users are quite against this phrase. They urge the users to re-think their negative attitude and aid the asker to address the issue by:

- Narrowing down the scope of the question
- Clarifying what they are asking
- Adding more details
- Describing what they have tried and show their code

1. Argumentation: This category was evident in the same rate on all 3 categories. The majority of the comments belonging in the *argumentation* category, included text that contradicted previous proposed answers or related issues in the comment section. When more than 1 proposed solutions were given in

the comment section, especially in evident in HRQ and HRUQ, users would argue against that proposal either by giving a new solution, perspective or would indicate why the suggested solution would not work.

2. Argumentation against the formulation of the question: Additionally, few comments could not be categorized as pure argumentation, as they aimed directly to the questions provided by the asker. This phenomenon did not show up in the HRQ and HRUQ as in LRQ. Related to the score of the questions, HRQ and HRUQ, had a solid foundation regarding the formulation of the questions, thus users did not argue against it. In comparison, the majority of the argumentation found in the LRQ "attack" the user's attitude (no effort on solving the issue on their own) along with the effort that they put in the formulation of the question. Thus, the category proved to be the second most prominent category in LRQ.

3. Inquiry for further explanation in a neutral manner: One of the most profound differences observed between HRQ and both LRQ/HRUQ. This category proved to be the 3rd most frequent for both LRQ and HRUQ. Quite recurrent was the phenomenon for HRUQ, that users would request further clarification regarding the issue at hand. users were trying to acquire as much information as possible around the issue of the user which many times resulted in the provision of external links and further discussion. The users would dedicate a plethora of comments trying to aid the user. In contrast, LRQ showed that most of the inquiry for information was done in relation to the argumentation against the formulation of the question. Users had difficulty to understand what the asker was trying to achieve, and especially what has been tried in the past, in order to be able to provide sufficient solutions.

4. Inquiry for further explanation in a positive manner: This category proved to be less populated in an equal rate for all groups. As it has been mentioned before SO aims in neutral discussions and despite the high rate of *common goals* (analysed below) for HRUQ, users were still more confident to require additional data in a neutral than a friendly manner. Few users of HRQ used, emojis while askers of LRQ required explanation to potential solutions or inquiries posted by other users, with text that showed gratitude.

Code	Definition	Text language examples	*HRQ*	*LRQ*	*HRUQ*
1. Argumentation	When the asker or another user, argues against a proposal made in a previous comment by another user/asker.	"your printm() is wonderful but…" "This is not a dupe as my issue focuses mostly on..." "Is this really a duplicate?" "This, does not answer my question"	18	16	15
2.Argumentation against formulation	When other users criticize the code, main body formulation, title, explanatory text, what have been tried so far, as posted by the original asker	"You should switch to Python 3.6 while you are still learning", "why have you not even attempted it?"	1	32	5
3. Inquiry for further explanation in a neutral manner	When another user, is neutrally asking for further clarification to the original question, or further evaluation to the methods, provided in the comments	It would be nice if you could provide further code, thank you!', "Have you tried setting the TensorFlow…report the results?"	6	28	30
4. Inquiry for further explanation in a positive manner	When another user, is positively asking for further clarification to the original question, or further evaluation to the methods, provided in the comments	Please provide us with the…. Thanks", "What OS are you using? :)"	3	5	2
5. Discussing the question in a neutral manner	Discussion around the solution to the issue at hand with neutral linguistic expressions, inside the SO guide norms	This is just a guess but...', 'Have you tried...', 'It would be possible to...'	41	48	94
6. Discussing the question in a more negative manner	Discussion around the solution to the issue at hand with negative linguistic expressions, outside the SO guide norms	"Have you even tried to read the description?" "How is someone supposed to help with no feedback on..." "Have you even tried to solve it on your own?"	1	4	0
7. Discussing the question in a friendly manner	Discussion around the solution to the issue at hand with friendly linguistic expressions, outside the SO guide norms jokes, funny expressions, "puns", emojis are also included	"Thank you for the heads up." "your printm() is wonderful", "Or the same if the computer crashes before :D", "finally will not execute if the power cord is ripped from the wall."	15	10	31
8. Common goals	Comments from users that are facing or have faced similar issues in the past, with possible solutions	"Any solution to this? I am facing the same issue" "Yep, same problem"	9	0	13
9a. Additional information	Comments that provide links to external resources, in an effort to provide additional information	You could take a look at this…' "There is a (github) that might solve your problem"	7	4	28
9b. Additional information (Site specialization)	Comments with links to internal resources, either in SO or other Stack platforms	"Could this question solve your issue?" "You might want to take a look at this"	13	18	9
10. Reminder of SO principles	Comments that have as a common theme to promote compliance to the SO principles this could include comments that only mention "this is a dupe/lacks clarity/ not focused" without further explanation as to why.	"This is a duplicate [link]", "The description lacks clarity" "This is not a coding service..." "Please refer to the S.O guide… before posting"	4	27	5

Table 6: Conversational Framework for the HRQ

5. Discussing the question in a neutral manner: The most evident category of the comment section in SO. It should be noted that the most significant differences between discussion and argumentation was the main topic of the comment. Comments marked as discussion-related did not have text that was "against" anything previously said or information provided by the asker. The most common themes were the provision of new solutions, ideas, or build-up in previous suggestions. Almost 100% increase in comments was shown in HRUQ in comparison to the other two groups. One potential reason behind this, was the constant participation of original askers in the comments, discussing with each

independent user about their issue, being eager to find a solution. Asker participation was evident in HRQ but not in the same level, as the asker would usually write fewer comments addressing to everything said before, while in LRQ, the majority of the discussions was done between users providing solutions.

6. Discussing the question in a more negative manner: The least observable pattern in the SO comment section. Even though it was hard to define it into a framework, the argumentation in the LRQ was more tense. Users turned to be more abrupt and passive-aggressive to the asker, especially when they have not even tried to solve the issue on their own, as discussed previously. But it was not considered in the level personally attacking the user. One such case was observed in the HRQ, but other users urged to support the defendant.

7. *Discussing the question in a friendly manner*: The second most evident category for HRUQ and quite high percentage in the rest groups, were comments that had a friendlier vibe. In HRQ, those comments mostly came by the high reputation askers. This showed that using a friendly tone and emojis in the comments is not seen as against the SO principles from the top Python reputable users. Another proof that the comment section is completely independent by the "main" platform and users have created their own rules in that section. The tones were not as informal and friendly as it would be in a small group chat between individuals, but at the same time it did remind some of the Reddit posts. Overall, users tended to use emojis to express themselves. Sad faces or any emojis that showed discomfort were used by users that face the same issue or the troubleshoot suggestions have not led to the required results. At the same time, some users might use happy faces and other emojis at times when they suggested a possible solution for the asker to test it or during argumentation. There were few examples where they used "puns" or expressions that were either hard to be recognized outsides memes or in a text-based format. In relation to LRQ, the community seemed to be more friendly and helpful to users that have tried in some way to solve the issue. Not only because it reduced the potential for wrong answers based on the lack of clarity but because it showed that some effort was put by the asker before they went online to ask for help. The amount of friendly discussions in HRUQ, even though it did not surpass the neutral tone discussions (see Table 6), was again more evident. The reason behind this, was that users of the community that faced the same issue tended to create a friendlier bond with the asker, discussing about the things that they tried so far and if they were using the same kind of hardware/software. At the same time, some askers that have sacrificed bounty to attract new answerers, feel keener to new users asking for further clarifications, thus they tended to be friendlier with them.

8. Common goals: Additionally, there were users that have faced the same issue in the past or are facing the same issue now. The latter group of users, described what they have tried so far (maybe some of those ideas could solve the issue) while the former, provided an answer but since it might not lead to the solution of the problem, they preferred to post it in the comment section (there might be a fear for possible downvotes if the answer is not corresponding to the issue). HRQ saw most examples of the former, where previously struggling users would provide their solutions. The users that faced currently common issues are a lot more profound in the unanswered questions. This was quite intriguing as with that percentage of users participating in the comment section, it could be argued that those questions would be solved immediately, as they were able to attract more potential answerers. In some cases, the question attracted new users, and created new discussions, months after the original post. In those cases, it could be seen as proof that the asker has chosen wisely the initial title, as it aided fellow users to NOT create a duplicate question which could potentially be closed as duplicate. Other users would suggest what they have also tried so far in the comment section (things that might

have stayed unnoticed to the asker). This led to the high number of friendly discussions. Additionally, a big percentage of the HRUQ, according to the comments, are unsolvable. As reported by the users, the issue at hand is unrelated to the code provided by the asker, but lied within the platform/software that the code belongs into. Unfortunately, no users admitted common goals in the LRQ's comment section.

9a/9b Additional information: Some users could be described as the SO info-savvy who would either provide similar discussion posts made in stack exchange, or other previous questions that have a solution that could potentially aid the user to find the solution that they were after. Those questions were following again a neutral tone, while providing links to SO or other stack platforms (see Table 6). The users have potentially established some sort of *mastery* around the platform as they were able to dig up information from the site and aid the asker to their issue. An interesting phenomenon, specific to the platform, which showed that the users were not eager to copy an already given answer to the problem and post it as their own, as it could lead to an unrelated solution to the initial problem of the user. Those type of users were mostly evident in HRQ. When external links were provided by the other users in LRQ, the majority of links focused in the Python dictionary and other educative websites which would provide the user with overall knowledge around the language but not specifically to their issue. The majority of external links shared to the HRUQ comment section were related to files of similar code, previous successful solutions to similar matters, among others. The most popular link used, once again, was GitHub with python "dictionaries" close second. The share of previously created SO threads was not as often observed.

10. Reminder of SO principles: Fewer times were when users would post a similar question just to mark the asker's as a duplicate (this could be related to a reputation/badge hunt). Those questions were grouped as *Reminder of SO principles* and NOT as additional information. The difference between those questions and the latter, was that the users did not put any effort to establish as to why they suggest the provided link. They usually created single sentences as "This is a possible duplicate of…" followed by the link provision. This kind of pattern mostly showed in LRQ. Most of those users were aiming to attract editors that could potentially close the LRQ, as a consequence of lack in clarity, lack of information, or as a duplicate. Usually, when the users referred to the SO principles in LRQ, they would additionally provide the links to SO guidelines and FAQ for the user to obtain the information that they needed around the correct formulation of a question.

Overall, it is important to note that comments seem to be not attracting the same attention as answers for the rest users of the community. Even though users can contribute by voting for "helpful" comment, in a similar way they vote for the answers they tend to not participate into the voting that much. This could be either that the comments were being made a long after the thread has been closed or solved or because the users were not that interested. This could potentially require further research. Additionally, some of the comments followed the "necromancy" style (see Figure 2). Users who might face the same issue but slightly varying and the best answer previously selected does not suit their needs, might resurrect the thread to attract again users to obtain a different answer. Sometimes users might create a whole conversation beneath the question in the format of the comment. This led to comments that request for a closed question to be reopened. A phenomenon occurring not only to old, closed threads but to recently ones as well.

At the same time, in regard to LRQ, if the user tried to introduce themselves as beginners in the comments, it did not make a difference in the downvote/upvote output, nor did the community show significantly "more" understanding. In some cases, users were able to see through the question and understand that it was related to an exercise. Even after providing a possible solution, they were trying to make known to the user that it is something that they should be learned through *Learning* and not

through SO. One extreme case that seemed quite significant, to shape an image around the community, was that the users of Stack have created or believe in certain ethics that are not presented by the pre-defined community guidelines. As for example, when one user tried to "scrape" info from a website that was blocking such behaviour, many users urged to consult the asker that this is not a good behaviour and they should rethink as it is unethical.

To conclude, it was observed that there is a relation between the user's reputation score and question score. The community seemed to support the SO guidelines and the upvote/downvote of a question was closely related to if the asker was formulating the question in relation to the framework of rules. The reputable members of SO seemed to be able to distinguish the correct formulation of an initial question (heading), along with following the SO principles regarding the utilization of a code snippet, a well formulated explanatory text, mentioned what they had done so far as well as through the usage of a clearer distribution when the question required lots of analysis. The majority of askers avoided the usage of informal language, while they utilized as many tags can be related to their issue. External resources were evident mostly in HRQ, informing the community as to the material that the asker had already analysed, either inside the SO platform, or with external links. The biggest discrepancy of HRUQ when compared to HRQ was related to the initial question. Even though the reputation of the users was high enough for them to be familiar with the platform (median of 570+ rating) users were potentially struggling with the correct utilization of a heading. More discrepancies were observed in relation to LRQ. The median reputation of the askers was considered significantly lower in comparison to the other groups. This subsequently showed that either the members of community were new, in regard to their contribution or new programmers in general. This was evident especially from being unable to comply to the SO guideline.

Additionally, the community utilized the comment section as a troubleshoot chat. Users tended to create discussions, by utilizing some patterns at the same level. Mostly, by *arguing* against solutions given by previous members *discuss in a neutral manner* without *expressing overall feelings*, while *inquiring, in a neural manner, for additional information* in regard to the asker or in regard to potential solutions shared from other users. Moreover, they seemed eager to provide a solution, and did not just discard a bad-formulated question, thus they commented to gain further clarification to the asker's issue. the discussion did not show signs of negative attitude towards the user (trolling or harassing) but they were likely to be passive-aggressive if there was no effort by the asker to solve the issue on their own. This was mostly evident through the common usage of the *"Stack is not a coding service"* comment which showed up in many of the LRQ. There were few examples of users that seemed to recognize the significance that low-skilled coding related issues should be solved through learning and not utilize the platform as a hand-out solution website.

8. Discussion

Through the analysis of formulation of questions and the discussion in the comment section, it has been possible to observe how the askers interact and exchange information, as well as the opinion of other users around this information exchange. The following section is going to focus on the discussion of the results through a conjunction with the learning theories around community learning as mentioned in the theoretical framework. Accordingly, I am going to focus on the main points of the results and compare them to previous research studies, while investigating which patterns, if any, could be related to the communities of practice.

The results of this study are going to be discussed in relation to the research questions:

What defines a high ranked question, regarding the community guidelines, in relation to unanswered and low-rated questions?

How do the users, of different levels of contribution (reputation) interact and inform the exchange of information (asked questions) through the discussion/argumentation in the comment section?

Does (this procedure of) information exchange share any similarities to the characteristics that are evident in a Communities of Practice?

8.1 Formulation of questions in relation to community's guidelines

According to the findings of this study, users of SO do not only agree with the guidelines of the platform regarding a well-established answer(Calefato et al., 2015), but there is also a consensus regarding question formulation.

It has been observed that there is a relation between the user's reputation score and question score. It is highly likely that users, by their constant participation in the platform, create a more methodically correct questions according to the community's guidelines. This was visible through the unique categories created and the latent analysis of the patterns. This partly verifies the results of Calefato et al. (2018) that users with high expertise in the platform are able to create successful questions. Even though they suggest that the reputation score is not always representing the ability of the asker to create successful questions, this study's findings showed that there is a relation between reputation score and the upvotes. Additionally, HRQ are possible to gather a higher score (80+ score) in contrast to LRQ which are limited to around -10 Even though the "popularity" of the asker or the time window for the posted question were not examined, the level of recognition as described by Chua and Banerjee (2015) did relate to the high score of a question. Thus, the level of reputation does affect the engagement of the users as it has been previously found in the relation to the selection successful answers (Duijn, Kucera, & Bacchelli, 2015).

As previous studies have shown the questions of the user could differ regarding their needs. HRQs and HRUQs can be separated between informative and conversational questions. Even though the majority of the questions observed were related to real life issues requiring a solution (Allamanis & Sutton, 2013), or what could be described as worked examples (Plass et al., 2010), there were examples of concept questions as well as questions that were related to transitions from other languages (Allamanis & Sutton, 2013). All the questions that required a solution, have included a code snippet, an

explanatory text, have mentioned what has been tried so far and were defined by a clear distribution. Users seem to similarly utilize bold and italic to highlight important software or code features as well as used greyed text to indicate libraries, etc. Questions that were what described as concept questions did not include all of the SO guidelines as their type was more conversational. The majority of HRQ and HRQU did not use expressions of feelings or informal language while they utilize as many tags can be related to their issue. External resources were evident, informing the community as to the material that the asker had already analysed, either inside the SO platform, or with external links.

There were not many observable differences regarding the HRUQ that were further analysed. The biggest discrepancy when compared to HRQ is related to the initial question (heading). Previous studies have shown that questions with shorter title and description are able to gather more responses by community users (Chua & Banerjee, 2015). Even though there was not a deep analysis regarding the length of the initial title, HRUQ seem to have extensive titles, when compared to HRQ and the majority of them were created in the format of a title, not a question. Even though the reputation of the users was high enough for them to be familiar with the platform (median of 570+ rating) users were potentially struggling with the correct utilization of a heading. Calefato et al. (2018) showed that a short title did not seem as important as a concise writing style for a successful question, hence a deeper analysis is required to create conclusive results regarding the title. HRUQ question bodies followed closely the guidelines with, sometimes, questions surpassing the clarity of the HRQ. Similarly, they did not use expressions of feelings or informal language while askers utilized external resources though not as often as in HRQ. Additionally, the findings showed a difference in the utilization of tags. Even though Chua and Banerjee (2015) suggested that fewer tags tend to attract more answers, HRUQ were not utilizing language specific tags (e.g. Python 3.1) something that was visible to the HRQ. Additionally, very few HRQ did not utilize less than 4 tags, findings which tended to agree with Saha et al. (2013) and their suggestion for a tag automated system for the correct utilization of all 5 tags allowed by the platform.

More discrepancies were observed in relation to LRQ. The median reputation of the askers was considered significantly lower in comparison to the other groups. This subsequently showed that either the users were new in regard to their contribution, or new programmers in general. This was evident especially from being unable to comply to the SO guideline. The platform does provide in the help section a “how to ask” sub-section where it is presented a way to formulate a good question, especially to new askers. According to the research of Calefato et al. (2018), the expression of feelings can dramatically affect the score of a question. Almost half of the LRQ askers of this study, showed some form of gratitude, distress. The edits showed that the users prefer a neutral style of writings, hence a big part of the editing was done in relation to the expression of feelings of the asker.

Given the fact that the majority of the LRQ were created from users new to programming and/or in Python in general, it seemed that this group of people in the majority, are individuals who were still learning the basics. According to Lave and Wenger (1991), users need a period to adapt to the norms of the new community by learning the form, expressions and the language used. Additionally, not every member of SO wants to be part of a community and they periodically use the site for info and feedback on their work, users that utilize the platform as “lurkers” or ghost users. This will be discussed more under Limitations.

8.2 Information Exchange through voluntary contribution

The findings showed that through the utilization of the comment section, users of the platform were able to aid the users not only with suggestive solutions but in regard to the formulation of the initial question. Users tended to create discussions, by utilizing some patterns at the same level. Mostly, by *arguing against* solutions given by previous users *discuss in a neutral manner* without *expressing overall feelings*, while *inquiring, in a neural manner, for additional information* in regard to the asker or in regard to potential solutions shared from other users. By utilizing the comment section as a trouble shoot chat, users were able to shape suggestive solutions with the usage of past experiences, ideas and thoughts (Chen & Hung, 2010).

Similarly the platform of SO provided the users with the actions for this knowledge sharing to happen, as with tools for the re-usage of the shared knowledge which seemed crucial according to previous research (Ba et al., 2001; Gang & Ravichandran, 2014). Though the ability of voting, editing as well as the reputation system, the platform conveyed the impression that it provides sufficient motivational factors for the users to not be hesitant in sharing their knowledge by providing rewards or other tangible benefits (Gang & Ravichandran, 2014). Previous research has pointed out that the trust factor is of most importance for users to share their knowledge to virtual communities (Chen & Hung, 2010; Hsu et al., 2007). There were no observable personal recognizability or identification-based trust (Hsu et al., 2007) between the members, not even in the two highest reputable members. Yet, the social exchange theory (Emerson, 1976) concept is evident, as users are voluntarily participating in this information exchange with the hope that they will be treated similarly in their time of need, hinting that potentially the platform provides some trust with its reputation system.

The level of contribution differed between the three distinct question groups with LRQ and HRUQ receiving most of the comments**.** According to Li et al. (2012) a well thought and structured question provided more chances for answers than an ill-informed one. As thus users of the platform directly or indirectly guided the asker to recognize their mistakes, suggested where their question lacks in detail and how it could be improved further through *argumentation against formulation* and *inquire for further explanation*. This same procedure was previously observed in the answer section where users aided the questioner by reformulating their question, even if that was not the main focus of the asker (Nasehi et al., 2012). Those patterns showed that through commenting, users were able to evaluate the quality of the content shared (Matschke et al., 2014) and reshape it to the community guidelines.

Previous research has showed that comments could provide significant changes to the flawed posted answers (Soni & Nadi, 2019). The findings seem to indicate the same concept for the creation of questions as users seem to take into consideration suggestions and edit their questions in order to attract more users. This proved to be especially visible in regard to HRUQ where the asker would edit their questions multiple times according to the discussion created in the comment section. This dedication of the HRUQ askers seemed to contradict the previous study's findings regarding answers, which proved that most of the comments are never noticed by the answerer, thus flawed answers remain flawed (Soni & Nadi, 2019). The HRUQ askers showed even more determination by actively participating in the comment section which in result created the double amount of *discussions around the issue in a neutral manner.* As thus, those results showed that self-efficacy is quite important in accordance to previous studies (Chen & Hung, 2010) where through the personal actions of the asker, they would increase the chances for them to acquire the knowledge needed. The high rating of those questions related to the findings of Sin et al. (2016) which showed that the active participation of the asker could relate in higher rating scores. Against the high number of recommendations received, the high number of participants as well as the diversity of the comments beneath the questions, HRUQ were unable to receive any answers. This has probably suggested that an "attractive" question body does not always provides answers as previous studies have shown (Adaji & Vassileva, 2016; Soni & Nadi, 2019).

Chua and Banerjee (2015) showed that the users of the platform were willing to aid new users, which was visible in the findings regarding the comment section. Users seemed eager to provide a solution, and did not just discard a bad- formulated questions (LRQ), thus they commented to gain further clarification to the asker's issue. Surprisingly, the discussion did not show significant signs of negative attitude towards the user (trolling or harassing) but they were likely to be passive-aggressive against the asker's effort to solve the issue on their own. In the few cases where the commenter's tone was too aggressive, even though they yielded high expertise, other users tended to support the asker and argue against the harasser. As thus it could be hypothesized that these type of comments can negatively affect the reputation of the commenter/answerer as previous research has showed in similar CQAs (Kim, 2010).

Finally, both HRUQ and LRQ showed significantly more presence of external resources. In the case of the HRUQ those references were supportive material and additional information regarding to the askers issue as *hints of peripheral processing (Freeman & Spyridakis, 2004)*. In regard to LRQ the references showed an educative perspective. Fewer users were more methodical, by doing a time-consuming search to external links in learning platforms that could potentially provide a solution to the user through learning or to previous questions that prompted a similar answer. Hence, there were users that recognized the significance that low-skilled coding related issues should be solved through learning and not utilize the platform as a hand-out solution website as has been previously suggested (Nasehi et al., 2012).

8.3 Stack Overflow as a Community of Practice

According to Preece and Maloney-Krichmar (2003) "there is no accepted definition of online community. The term means different things to different people." As a result, we cannot certainly say if a platform like SO could be seen as a CoP according to Lave and Wenger's (1998) but it shared most of the characteristics evident in any CoP. Furthermore, it is possible though to discuss the findings of the study in relation to the theory.

It has been considered that a community that yields the characteristics of a community of practice, should provide users with multiple viewpoint characteristics of participation (Fisher, Erdelez, & McKechnie, 2005; Radford et al., 2017), or what Wenger (1998) considers as *activities*. Indeed, SO provides the users with the chance to post and answer questions, vote positively or negatively, post comments while providing editing functions for the active users of the platform. Additionally, SO is open to everyone and any user can become a non-participant observant of it.

Previous research has showed the *shared repertoire* that users of SO are sharing, is filled with an unlimited amount of information (Vasilescu et al., 2014) while SO provide learners with ways to find the content that they seek (Vassileva, 2008). Except for the algorithm and the gamification features of SO, the community seems to play a significant role in shaping the *shared repertoire*. According to De Moor (2006); Rosenbaum and Shachaf (2010), this *shared repertoire* could include everything from the information previously created, tasks and discussions, along with the, potentially, accomplished goals. This study partially showed that this CoP characteristic was evident in SO. The collection of collaborative work is furthermore consisted of previously answered and unanswered questions, regardless of their final score, the interaction beneath it (comment section), along with the edits. According to the study's findings, even once the question received a successful answer, users were still editing question bodies, titles while making the overall post as concise and descriptive as possible so that future users would be able to attain the information that they seek. The level of editing varied greatly from fixing minor issues in the HRQ, to the creation of extensive text, additional code and reshape of the title in the LRQ. This procedure could be described as part of the *practice* that must be evident in a community for it to be considered a CoP (Rosenbaum & Shachaf, 2010; Wenger, 1998).

Users of the platform were able to create this abundance of information by sharing their professional experiences in order for the asker to attain knowledge and meaning to the solution, similarly to Kolb's experiential theory (Kolb, 1984). As thus the community is the *active contributor* of knowledge and what Wenger (1998) describes as *learning*.

The community did not partially relate to the guidelines when producing a question. According to the questions scores and the re-formulation of questions, users actively supported this *framework of rules* as it seems to represent the majority of the community. It is thus evident that the *framework of rules* has been shaped based on the needs of its users. This was visible during the analysis of this study and coincides with the Rosenbaum and Shachaf (2010) research that the users of those digital communities, constantly interact, constitute and re-constitute the social worlds in which they coexist. The findings showed that HRQ who followed all of the SO principles to have higher scores than those who were missing code snippets or *what has been tried in the past* (LRQ). According again to the edits, users seemed to focus on creating questions, as close as possible to the guidelines provided. Additionally, LRQ showed that users of the platform heavily criticized the formulation of the questions, while providing constant reminders to the asker of the SO principles. Through this procedure it has been evident that users foster a feeling of co-accomplishment through repetition and re-evaluation (Berlanga et al., 2008; Lee, 2004).

The *mutual engagement* and interaction of the users was visible through the comment analysis, where users engage in the pursue of a solution by sharing a *common goal*, coding. Through this *voluntarily* comment interaction and support, they were possibly able to create what Wenger (1998) describes as a *joint enterprise*. Through their constant active participation, common goal and guidelines they were able to show their expertise with their real-life past coding experiences. According to the analysis of comments, it was evident that the *community* is highly important for the *learning* to take place. The acquisition of a successful solution was not usually given irresponsibly without the asker having understood as to why it should be done this way. Another example of that could be seen when users were not directly giving a potential solution to LRQ askers and guide them to learn how to solve "simple" tasks by the provision of external links to educational material.

Additionally, some of the drawbacks of a CoP were observed. Through the creation of passive-aggressive comments against the formulation of the question and with the common usage of the "*Stack is not a coding service*", many users did not develop further in the contribution of information in the platform. This was indicated by their reputation score which still remains considerably low. This could be related to the insufficient time for the users to adapt in the community standards and personally develop as in any other CoP (Correia et al., 2010), which in result leads to trust issues, and low level of cooperation, creating even more non-participant users to the platform. Even though SO and its strict guidelines protect most users from bullying and harsh language this does not mean that less eminent comments could not hurt users of the community. Thus, comments such as "SO is not a coding service" and other similar passive aggressive attitude comments could harm those users, leading them to become ghost users of the platform.

To conclude, it seems that the platform of Stack Overflow could be considered a CoP as described by Wenger (1998), as the platform's guidelines and shared repertoire are highly dependent on the users' needs, while the site provides the users with the ability to mutually, voluntarily engage in the information exchange, who then are rewarded for their contribution in the creation of new knowledge through the identity acquisition system (reputation).

8.4 Limitations

One of the most profound limitations of this study is in regard to the method used. In order to any form of content analysis to be successful, the creation of categories, the explanation along with examples given should be created, used, analysed and revised by more than one researcher. Being a master's thesis, the coding happened only by one person as thus the categories created, especially in relation to the formulation of questions, might not be considered sufficient. Additionally, a researcher conducting content analysis, could fail to develop a complete understanding of the data which could result in key categories missing (Hsieh & Shannon, 2005). Thus, there was no dedicated effort in examining Krippendorff's alpha or COHENS kappa around the agreement and validity of the categorization as suggested by (Bourgonjon, Vandermeersche, De Wever, Soetaert, & Valcke, 2016). Bias can be a hindering factor in this type of analysis and data selection (selective, partial or biased); thus, it is most commonly conducted by more than one researcher, especially when the material is categorized manually without the usage of computer aided analysis.

Employing a mixed methods content analysis, methods from ethnographic observation were used, during the first read-through of the material gathered and before the creation of categories. However, a researcher might miss-read or miss-calculate the idea behind the creation of comments and questions as this type of data was not created with the researcher in mind (Hsieh & Shannon, 2005). Additionally, some words or patterns observed through the formulation of questions and comment analysis, might be included in multiple categories but could potentially have less significance in one of them (Cohen et al., 2011). Even though, in the case of the comment analysis, the coding framework was developed on top of another successful framework, this does not mean that the initial idea of those researchers (Haythornthwaite et al., 2018), corelates with this study's general aims. Thus, it required modifications based on pre-existing categories.

The quantitative results presented in this thesis, barely represented frequencies among the patterns created along with the usage of calculation of means and median for the users' and questions' score. A more advanced researcher would be able to successfully create descriptive statistics to successfully investigate the correlation and statistical discrepancies between the different groups of questions and the categories that are part of.

Additionally, as argued by Elo and Kyngäs (2008): "The analysis process and the results should be described in sufficient detail so that readers have a clear understanding of how the analysis was carried out and its strengths and limitations". At the same time the researcher is responsible when conducting content analysis, to thoroughly analyze the process in as much detail as possible. Even though, I tried to be specific in the creation of the categories along with the analysis of the patterns observed by providing the researcher's own actions and insights(Elo & Kyngäs, 2008)**,** it might not have been conducted in a sufficient manner to be easy for the reader to observe the distinct differences among the different categories. This has been recognized by other researchers who have argued that content analysis is far more complex and less standardized than a quantitative analysis(Elo & Kyngäs, 2008; Polit & Beck, 2004).

Additionally, it should be re-mentioned that I was a user of the site few months before I started the work on this thesis. This has potentially led to significant benefits, as I was familiar with the environment and I was able to delve deeper in the issues that new users are facing but at the same time it could have resulted in bias especially for the selection of patterns and analysis of those. This was one of the reasons as to the plethora of previous research that is available in the thesis, in an effort to increase the validity of my methods, results and my hypotheses.

Except for the general investigation regarding the formulation of questions and how users inquire for information in relation to the community's framework, an effort was put to see if SO attains

Community of Practice characteristics. This effort was based on previous conducted research along with a brief analysis of the CoPs as presented either by Wegner or other researchers.

Wenger (1998), suggests that users of such community "move towards full participation in the sociocultural practices of the community". This move is achieved by establishing relations between new and more experienced users and through the activities that they participate in. At the same time, it is recognised that CoP includes users that do not actively participate in those activities relying only on the observation of relations between the active users. Those users are present in CQAs. Even though, this study tried to investigate how the active users comment and move from a question to its potential solution, only a part of the active community was investigated. As thus the material analysed was based on questions and comments on those questions, while data was not collected regarding answers and comments to those answers. As a result, this study provides a partial view of the community. Additionally, there is not sufficient amount of research conducted in relation to the non-active users, most commonly described as "lurkers" or "ghost users" from researchers. This could be potentially related to how hard it is to investigate and gather material around this part of the community in order to create assumptions regarding to their involvement in a CQA or CoP in general. Accordingly, it could be hypothesized that there are different levels of non-participant users. The SO platform provides anonymity in some actions (such as voting, reporting) which do not result in any visible level of contribution for a researcher and could be a distinct category of inactive users, when compared to plain readers of posts.

8.5 Suggestion for future research

One suggestion for future research would be the revision of this study, with the aid of fellow researchers. Through the collaboration of different mindsets, it would be possible to de-limit the reliability and validity of the results, as well as conduct a deeper investigation regarding the correlation between the user's reputation score and question voting score.

Moreover, a survey could be conducted to investigate the users' opinion regarding the framework of rules presented by SO, as well as their thoughts around the importance of the comment section and a well-formulated question. It could also be investigated if the platform should include in its gamification features a "no upvoting/downvoting" concept to new contribution member posts as it hinders their confidence and possibly leaves them stagnated. This could potentially lead to the comment section to be utilized as an *advisor* regarding the SO principles and educate new users to the importance of the community standards.

As previously mentioned, the sample of this study was related to the question and the comment section beneath those questions. As a result, only a fragment of the community interaction was examined. It is advisable that the same research could be conducted to investigate the way that the users create and formulate their answers, along with the interaction in the comment section beneath those answers.

Furthermore, for a closer look in the interpersonal relations between the users, the comment interconnection and the patterns of dependency in a deeper level, a form of discourse analysis could be used in the comment section.

9. Conclusion

This study focused on the importance of a high-ranked question according the community standards. It additionally showed how the patterns of users' engagement in the comment section affect the information acquisition procedure and the creation of suggestive solutions. Additionally, the characteristics of this information exchange and the volunteered interaction of the users were compared to the characteristics that must be evident in a community, in order to be considered a CoP.

The formulation of a question is the epitome for the existence of a platform like SO. Through the creation of a successful question the asker is able to acquire a solution, acquire new knowledge but indirectly aid future programmers that might face the same issue. Furthermore, it is the first step for a new user to familiarize themselves with the community's framework of rules and begin their integration to the community. As thus, the creation of a successful question is critical, not only for the original poster but for the community as a whole. Through this study it was possible to create a raw idea as to how the high-rated questions differ in frequencies in relation to the low-rated ones and the minimal differences with the unanswered. It was also visible that the community users tend to agree and support the SO principles, proving that the question formulation framework is adapted to the needs of the users. Additionally, it became evident that there is some relation between the level of contribution of the users (reputation system) and the score of a question as the significant majority of LRQ was done from low reputation users. Even though SO provides specific guidelines not everyone is willing to read them extensively and become a long-time member of the community. Similarly, it has been observed that through the discussion in the comment section, community users argument against the question formulation, providing solutions as to how it could be fixed by inquiring information, providing external links and utilizing the edit function. A big part of the community does not see the platform as a hand-out solution provider, neither as a starting point for beginners. The users are interacting and collaborating through troubleshooting, by sharing past and current experiences, while the level of participation differs. Overall, they share the same *guidelines*, the same *repertoire* and create a *joint enterprise* while maintaining and evolving the platform based on their needs. All the above entwine the characteristics of Communities of Practice regarding the *mutual engagement, common goal, shared repertoire, joint enterprise* and the importance of *a framework of rules*.

Overall, it seems that the platform of SO could be considered as an online community of practice for advanced hobbyists or professional programmers and users that are willing to comply to the SO guidelines. Even though it is one of the most studied platforms, there are certainly research gaps in order to investigate and evaluate the knowledge-sharing characteristics (learning potential), the deeper understanding of interaction of its users as well as the importance of the platform in the lifelong development of the programming community.

Reference list

Adaji, I., & Vassileva, J. (2016). *Modelling user collaboration in social networks using edits and comments.* Paper presented at the Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization.

Allamanis, M., & Sutton, C. (2013). *Why, when, and what: analyzing stack overflow questions by topic, type, and code.* Paper presented at the 2013 10th Working Conference on Mining Software Repositories (MSR).

Althoff, T., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2014). How to Ask for a Favor: A Case Study on the Success of Altruistic Requests. *arXiv.org*.

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2013). *Steering user behavior with badges.* Paper presented at the Proceedings of the 22nd international conference on World Wide Web.

Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and barriers to participation in virtual knowledge-sharing communities of practice. *Journal of knowledge management*.

Asaduzzaman, M., Mashiyat, A. S., Roy, C. K., & Schneider, K. A. (2013). *Answering questions about unanswered questions of stack overflow.* Paper presented at the 2013 10th Working Conference on Mining Software Repositories (MSR).

Ba, S., Stallaert, J., & Whinston, A. B. (2001). Research commentary: introducing a third dimension in information systems design—the case for incentive alignment. *Information systems research, 12*(3), 225-239.

Baker-Eveleth, L., Sarker, S., & Eveleth, D. M. (2005). Formation of an Online Community of Practice: An Inductive Study Unearthing Key Elements. In (pp. 254b-254b).

Bates, A. T. (2018). *Teaching in a digital age: Guidelines for designing teaching and learning*.

Berlanga, A. J., Sloep, P. B., Kester, L., Brouns, F., Rosmalen, P., & Koper, R. (2008). Ad hoc transient communities: towards fostering knowledge sharing in learning networks. *International Journal of Learning Technology, 3*(4), 443-458.

Bourgonjon, J., Vandermeersche, G., De Wever, B., Soetaert, R., & Valcke, M. (2016). Players' perspectives on the positive impact of video games: A qualitative content analysis of online forum discussions. *new media & society, 18*(8), 1732-1749.

Calefato, F., Lanubile, F., Marasciulo, M. C., & Novielli, N. (2015). Mining Successful Answers in Stack Overflow. In (Vol. 2015-, pp. 430-433).

Calefato, F., Lanubile, F., & Novielli, N. (2018). How to ask for technical help? Evidence-based guidelines for writing questions on Stack Overflow. *Information and Software Technology, 94*, 186-207.

Chao, C.-Y., Hwu, S.-L., & Chang, C.-C. (2011). Supporting interaction among participants of online learning using the knowledge sharing concept. *Turkish Online Journal of Educational Technology-TOJET, 10*(4), 311-319.

Chen, C.-J., & Hung, S.-W. (2010). To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities. *Information & management, 47*(4), 226-236.

Chen, I. Y. (2007). The factors influencing members' continuance intentions in professional virtual communities—a longitudinal study. *Journal of Information Science, 33*(4), 451-467.

Chinn, P. L., & Kramer, M. K. (1983). Theory and nursing a systematic approach.

Chua, A. Y., & Banerjee, S. (2013). So fast so good: An analysis of answer quality and answer speed in community Q uestion-answering sites. *Journal of the American Society for Information Science and Technology, 64*(10), 2058-2068.

Chua, A. Y. K., & Banerjee, S. (2015). Answers or no answers: Studying question answerability in Stack Overflow. *Journal of Information Science, 41*(5), 720-731. doi:10.1177/0165551515590096

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education*. London, UNITED KINGDOM: Taylor & Francis Group.

Correia, A., Paulos, A., & Mesquita, A. (2010). Virtual Communities of Practice: Investigating Motivations and Constraints in the Processes of Knowledge Creation and Transfer. *Electronic Journal of Knowledge Management, 8*(1), 11-20.

Daniel, B. K., O'Brien, D., & Sarkar, A. (2007). User-centered design principles for online learning communities: A sociotechnical approach for the design of a distributed community of practice. In *User-centered design of online learning communities* (pp. 54-70): IGI Global.

De Moor, A. (2006). Community memory activation with collaboration patterns.

Deterding, S. (2012). Gamification: designing for motivation. *interactions, 19*(4), 14-17.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *CHI'11 extended abstracts on human factors in computing systems* (pp. 2425-2428).

Downe-Wamboldt, B. (1992). Content analysis: method, applications, and issues. *Health care for women international, 13*(3), 313-321.

Duijn, M., Kucera, A., & Bacchelli, A. (2015). *Quality questions need quality code: Classifying code fragments on stack overflow.* Paper presented at the 2015 IEEE/ACM 12th Working Conference on Mining Software Repositories.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115.

Emerson, R. M. (1976). Social exchange theory. *Annual review of sociology, 2*(1), 335-362.

Eraut*, M. (2004). Informal learning in the workplace. *Studies in continuing education, 26*(2), 247-273.

Ferguson, R., Wei, Z., He, Y., & Buckingham Shum, S. (2013). *An evaluation of learning analytics to identify exploratory dialogue in online discussions.* Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge.

Fisher, K. E., Erdelez, S., & McKechnie, L. E. (2005). *Theories of information behavior*: Information Today, Inc.

Fogg, B. J. (2003). *Prominence-interpretation theory: Explaining how people assess credibility online.* Paper presented at the CHI'03 extended abstracts on human factors in computing systems.

Freeman, K. S., & Spyridakis, J. H. (2004). An examination of factors that affect the credibility of online health information. *Technical Communication, 51*(2), 239-263.

Gang, K., & Ravichandran, T. (2014). Exploring the determinants of knowledge exchange in virtual communities. *IEEE Transactions on Engineering Management, 62*(1), 89-99.

Gazan, R. (2006). Specialists and synthesists in a question answering community. In (Vol. 43).

Gray, B. (2005). Informal learning in an online community of practice. *International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, 19*(1).

Harper, F. M., Moy, D., & Konstan, J. A. (2009). Facts or friends? distinguishing informational and conversational questions in social Q and A sites. In (pp. 759-768).

Harper, F. M., Moy, D., & Konstan, J. A. (2009). *Facts or friends? Distinguishing informational and conversational questions in social Q&A sites.* Paper presented at the Proceedings of the sigchi conference on human factors in computing systems.

Harwood, T. G., & Garry, T. (2003). An overview of content analysis. *The marketing review, 3*(4), 479-498.

Haythornthwaite, C., Kumar, P., Gruzd, A., Gilbert, S., Esteve del Valle, M., & Paulin, D. (2018). Learning in the wild: coding for learning and practice on Reddit. *Learning, Media and Technology, 43*(3), 219-235. doi:10.1080/17439884.2018.1498356

Hsieh, G., Kraut, R. E., & Hudson, S. E. (2010). *Why pay? Exploring how financial incentives are used for question & answer.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288.

Hsu, C.-L., & Lin, J. C.-C. (2008). Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation. *Information & management, 45*(1), 65-74.

Hsu, M.-H., Ju, T. L., Yen, C.-H., & Chang, C.-M. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. *International journal of human-computer studies, 65*(2), 153-169.

Johnson, S. J., Blackman, D. A., & Buick, F. (2018). The 70: 20: 10 framework and the transfer of learning. *Human Resource Development Quarterly, 29*(4), 383-402.

Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. *Educational psychologist, 50*(4), 313-334.

Kim, S. (2010). Questioners' credibility judgments of answers in a social question and answer site. *Information Research, 15*(2), <xocs:firstpage xmlns:xocs=""/>.

Kolb, D. (1984). Experiential Learning, Englewood cliffs: Ed. In: Prentice Hall.

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *The International Review of Research in Open and Distributed Learning, 12*(7), 74-93.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*: Sage publications.

Lakhani, K. R., & Von Hippel, E. (2004). How open source software works:"free" user-to-user assistance. In *Produktentwicklung mit virtuellen Communities* (pp. 303-339): Springer.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*: Cambridge university press.

Lee, D. (2004). *Sense of co-accomplishment in collaborative work as threshold in establishing a sense of community in an online course.*

Lee, G. K., & Cole, R. E. (2003). From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organization science, 14*(6), 633-649.

Li, B., Jin, T., Lyu, M. R., King, I., & Mak, B. (2012). *Analyzing and predicting question quality in community question answering services.* Paper presented at the Proceedings of the 21st International Conference on World Wide Web.

Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G., & Hartmann, B. (2011). *Design lessons from the fastest q&a site in the west.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Manuti, A., Pastore, S., Scardigno, A. F., Giancaspro, M. L., & Morciano, D. (2015). Formal and informal learning in the workplace: a research review. *International journal of training and development, 19*(1), 1-17.

Marsick, V. J., & Watkins, K. E. (2001). Informal and incidental learning. *New directions for adult and continuing education, 2001*(89), 25-34.

Matschke, C., Moskaliuk, J., Bokhorst, F., Schümmer, T., & Cress, U. (2014). Motivational factors of information exchange in social information spaces. *Computers in Human Behavior, 36*, 549-558.

Nasehi, S. M., Sillito, J., Maurer, F., & Burns, C. (2012). *What makes a good code example?: A study of programming Q&A in StackOverflow.* Paper presented at the 2012 28th IEEE International Conference on Software Maintenance (ICSM).

Park, J., Konana, P., Gu, B., Leung, A. C. M., & Chung, A. (2010). *An Investigation of Information Sharing and seeking Behaviors in Virtual Communities.* Paper presented at the ICIS.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*: SAGE Publications, inc.

Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge

New York: Cambridge New York : Cambridge University Press.

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods*: Lippincott Williams & Wilkins.

Ponti, M. (2015). "Remember to hand out medals": Peer rating and expertise in a question-and-answer study group. *International Review of Research in Open and Distributed Learning, 16*(2), 327-350.

Preece, J., & Maloney-Krichmar, D. (2003). Online communities: focusing on sociability and usability. *Handbook of human-computer interaction*, 596-620.

Radford, M. L., Connaway, L. S., Mikitish, S., Alpert, M., Shah, C., & Cooke, N. A. (2017). Shared values, new vision: Collaboration and communities of practice in virtual reference and SQA. *Journal of the Association for Information Science and Technology, 68*(2), 438-449. doi:10.1002/asi.23668

Regner, T. (2014). Social preferences? Google Answers. *Games and Economic Behavior, 85*(C), 188-209. doi:10.1016/j.geb.2014.01.013

Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society for Information Science and Technology, 61*(9), 1933-1944. doi:10.1002/asi.21340

Saha, A. K., Saha, R. K., & Schneider, K. A. (2013). *A discriminative model approach for suggesting tags automatically for stack overflow questions.* Paper presented at the 2013 10th Working Conference on Mining Software Repositories (MSR).

Seaman, C. B. (2002). The information gathering strategies of software maintainers. In (pp. 141-149).

Self, C. S. (1996). An integrated approach to communication theory and research. *(Credibility. In M. Salwen & D. Stacks Eds)*(Mahwah, NJ: Lawrence Erlbaum Associates.).

Shachaf, P. (2009). The paradox of expertise: is the Wikipedia Reference Desk as good as your library? *Journal of Documentation, 65*(6), 977-996. doi:10.1108/00220410910998951

Shah, C., Oh, J., & Oh, S. (2008). Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday, 13(9)* doi:https://doi.org/10.5210/fm.v13i9.2182

Sin, S. C. J., Lee, C. S., & Theng, Y. L. (2016). Social Q&A question-and-comments interactions and outcomes: A social sequence analysis. In (Vol. 10075, pp. 325-338).

Singh, V., Twidale, M. B., & Nichols, D. M. (2009). *Users of open source software-how do they get help?* Paper presented at the 2009 42nd Hawaii International Conference on System Sciences.

Skinner, B. F. (1968). *The technology of teaching*: New York : Appleton-Century-Crofts.

Smith, P., Barty, K., & Stacey, E. (2005). *Limitations of an established community of practice in developing online innovation.* Paper presented at the ODLAA 2005: Breaking down boundaries: international experience in open, distance and flexible education. Proceedings of the 17th ODLAA conference.

Soni, A., & Nadi, S. (2019). Analyzing comment-induced updates on stack overflow. In (Vol. 2019-, pp. 220-224).

Sonnentag, S., Niessen, C., & Volmer, J. (2006). Expertise in software design. In.

Sowe, S. K., Stamelos, I., & Angelis, L. (2008). Understanding knowledge sharing activities in free/open source software projects: An empirical study. *The Journal of Systems & Software, 81*(3), 431-446. doi:10.1016/j.jss.2007.03.086

Srba, I., & Bielikova, M. (2016). Why is stack overflow failing? preserving sustainability in community question answering. *IEEE Software, 33*(4), 80-89.

Su, Q., Pavlov, D., Chow, J. H., & Baker, W. C. (2007). Internet-scale collection of human-reviewed data. In (pp. 231-240).

Thompson, T. L. (2011). Work-learning in informal online communities: evolving spaces. *Information Technology & People*.

Wang, S., Chen, T.-H. P., & Hassan, A. E. (2018). How do users revise answers on technical Q&A websites? A case study on Stack Overflow. *IEEE Transactions on Software Engineering*.

Vasilescu, B., Serebrenik, A., Devanbu, P., & Filkov, V. (2014). *How social Q&A sites are changing knowledge sharing in open source software communities.* Paper presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing.

Vassileva, J. (2008). Toward social learning environments. *IEEE transactions on learning technologies, 1*(4), 199-214.

Weber, R. P. (1990). *Basic content analysis*: Sage.

Wenger, E. (1998). *Communities of practice : learning, meaning, and identity*: Cambridge : Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. M. (2002). It takes a community.(excerpt from Cultivating Communities of Practice: A Guide to Managing Knowledge)(Excerpt). *CIO, 15*(15).

Vygotsky, L. (1964). Thought and language. *Bulletin of the Orton Society, 14*(1), 97-98. doi:10.1007/BF02928399

Yang, L., Bao, S., Lin, Q., Wu, X., Han, D., Su, Z., & Yu, Y. (2011). *Analyzing and predicting not-answered questions in community-based question answering services.* Paper presented at the Twenty-Fifth AAAI Conference on Artificial Intelligence.

Yardley, S., Teunissen, P. W., & Dornan, T. (2012). Experiential learning: transforming theory into practice. *Medical teacher, 34*(2), 161-164.

Zhang, H., Wang, S., Chen, T.-H., & Hassan, A. E. (2019). Reading Answers on Stack Overflow: Not Enough! *IEEE Transactions on Software Engineering*.

Zheng, Y., Zhao, K., & Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision support systems, 56*, 513-524.

Ziegler, M. F., Paulus, T., & Woodside, M. (2014). Understanding informal group learning in online communities through discourse analysis. *Adult Education Quarterly, 64*(1), 60-78.