



OULUN YLIOPISTO
UNIVERSITY of OULU

Evaluating the Knowledge Acquired from Online Q&A Sites: An Exploratory Study

University of Oulu
Faculty of Information Technology and
Electrical Engineering
Master's Thesis
Yini Wang
17.04.2018

Abstract

When people encounter questions, they often search for answers/solutions from Internet nowadays. Online question and answer (Q&A) sites are helpful for people to seek answers to their questions. On online Q&A sites (such as Stack Overflow), millions of questions have been answered. Sometimes, even one question receives dozens of answers. Evaluation of the knowledge acquired through the answers becomes an important but challenging task for knowledge seekers. This process of knowledge evaluation can be comprehensively understood by revealing all factors that knowledge seekers take into consideration when they evaluate the answers on online Q&A sites. Investigating the knowledge evaluation process on online Q&A sites can not only guide knowledge seekers to acquire knowledge more efficiently, but also shed light on managing the community of online Q&A sites to make them more user friendly.

A qualitative research was conducted, in which sixty-seven participants of a large online Q&A sites (i.e., Stack Overflow) were interviewed. The results of data analysis reveal that answer-indicators, living example, lengthiness, empirism, feasibility, readability, source credibility, and recency are important when people evaluate knowledge on online Q&A sites. A dual process is found to be employed by individuals during the knowledge evaluation process on online Q&A sites, in which they assess not only the answer content itself, but also heavily rely on the heuristic cues of the answer. Factors generated through this research also correspond to prior research on information quality.

Keywords

knowledge evaluation, dual process theory, information quality, online Q&A, Stack Overflow

Supervisor

Assistant Professor, University of Oulu, Li Zhao

Foreword

After a few months of hard working, today is the day for me to proudly announce the accomplishment of my Master's Thesis.

The duration of writing Master's Thesis is a process of deriving knowledge and improving myself. The shape of scientific world becomes clearer to me after I finished the journey of writing thesis. Luckily, I was not alone on this journey. Many people have helped me throughout the way. Thus, I would like to bring my sincerely appreciation to those generous and helpful people.

First of all, I would like to thank my thesis supervisor Asst. Professor Mr. Li Zhao at University of Oulu. You brought me to the scientific world. Then guided me all along when I was writing my thesis. You are always very patient to answer every single question from me and helpful with proposing practical advices regarding my thesis. My thesis has been promoted a lot with your guidance.

Next, I would like to express my thanks to Post Doc. Researcher Mrs. Dorina Rajanen at University of Oulu for her helpful suggestions on enhancing my thesis. You know that my graduation date is getting close, so you read through my thesis on your own time and gave me feedback promptly. I am extremely appreciated for your understanding and for your useful advices.

Then, I definitely should bring my thankfulness to those people who joined the interview for this research. I thank you all for your time and your valuable information provided. Without your help, this research would not be achievable.

I would also like to thank my dear friends, especially those who have spent plenty of time with me in the library from day to night. We discussed a lot on scientific writing and approaches for scientific research. You inspired me sometimes. We helped each other all the way. I wish you all best luck on your thesis work.

In the end, I need to thank my beloved families, who are always there for me. Now I can finally say to my dear grandma that I have done with my Master's Thesis when she asks me about the progress!

All in all, thank you so much everyone!

Yini Wang

Oulu, April 17, 2018

Contents

Abstract	2
Foreword	3
Contents	4
1. Introduction	5
2. Theoretical Background and Literature Review	7
2.1 Dual process theory	7
2.1.1 Central route: information quality theory	9
2.1.2 Peripheral route: intuitive processing	12
3. Methodology	14
3.1 Research site	14
3.2 Data collection	15
3.3 Interviews.....	16
4. Data analysis and results	19
4.1 Main factors that affect individuals' knowledge evaluation process.....	22
4.2 Popularity of main factors that affect knowledge evaluation process	24
5. Interpretation Based on Theories	26
5.1 Interpretation based on dual process theory	26
5.2 Interpretation based on information quality literatures.....	31
6. Discussion	35
6.1 Discussion of findings	35
6.2 Contributions to the literature	35
6.3 Implications for practice	36
6.4 Limitations and future research	36
7. Conclusion.....	38
References	39

1. Introduction

Nowadays Internet holds enormous amount of information, which is accessible for Internet users (Brinkley, & Burke, 1995). Many tools have been developed for making it easy for users to retrieve information from Internet. Online question and answer (Q&A) site is one of those. Online Q&A sites are commonly used as knowledge repository, where users can find potential answers to their questions (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2012). Some online Q&A sites do not have a specific topic, so that the discussion can emerge from various areas. Moreover, there are also online Q&A sites which constraint a specific topic like Stack Overflow.

Stack Overflow is an online Q&A site for programmers and developers. Most discussions on that Q&A site are about programming. Site users can ask questions there, while others can provide answers to the questions they are interested. According to the statistics found on Stack Overflow, there are over fifty million professional programmers visit the community monthly (Stack Overflow, 2018.). Stack Overflow is even used as the unofficial programming documentation by some developers (Gyongyi, Koutrika, Pedersen, & Garcia-Molina, 2007). It can be clearly seen that Stack Overflow is a popular online Q&A site among developers.

While Internet brings stupendous benefits to its users, it also brings obstacles that could embarrass its users in retrieving knowledge. One main obstacle is that the quality of answers on such Q&A sites is challenging to be evaluated since it varies from high-quality to low-quality, and sometimes it could even be abuse and spam (Agichtein, Castillo, Donato, Gionis, & Mishne, 2008). In addition, the answer might be given by some anonymous users from the community, from newbie to experts. That makes it painful for site users to check the source credibility. Furthermore, Internet environment inhibits the normal social interaction between humans on one level, which is due to the asynchronism and geographically distribution. The result of weakening normal social interaction is that a user loses some clues that are helpful for judging the source credibility. Yet, those clues might be received from normal social communication (Fadel, Meservy, & Jensen, 2015.). All of the various barriers listed above add more difficulties and complexity to the knowledge evaluation process. It is helpful to understand the knowledge evaluation process by conducting the research which can reveal the factors that individuals take into consideration when evaluating knowledge on online Q&A sites. This type of research would be useful for eliminating difficulties and barriers for individuals' knowledge evaluation process over Internet.

Figuring out how individuals conduct evaluation behavior on online Q&A sites is significant to information filtering literatures. Even though there is existing work (Fadel et al., 2015; Nasehi, Sillito, Maurer, & Burns, 2012) regarding similar topic. To our knowledge, research that investigates people's evaluation process for adopting answer on online Q&A sites still has much space to be fulfilled. Thus, this research can enrich the information filtering literatures by addressing an unfully-explored field of relevance.

Moreover, understanding the factors that knowledge seekers concern can light up a hint for online Q&A sites community, about how to write a high-quality answer that will likely be adopted by other users. With the same knowledge, online Q&A sites administrators would be aware of how to adjust the layout and content on the website, so that it will make knowledge evaluation process smoother and more efficient.

In this research, a qualitative study was conducted. The research work is explained as following. Firstly, prior literatures that are relevant to this research are reviewed. Then, the research methodology used in this study is introduced. After that, data that are retrieved from this research is analyzed and findings are presented. Next, the findings are interpreted based on relevant theories and literatures. After then, the results of this research are discussed. In addition, contributions of this research to the literatures are raised. Meanwhile, implications for practice are presented, limitations are discussed, and future work is suggested. At the end, the conclusion is asserted.

2. Theoretical Background and Literature Review

Evaluating knowledge on online Q&A sites embeds a filtering process, wherein individuals filter among several alternatives. The filtering criteria may vary depending on different types of knowledge seekers. After the filtering process, a satisfactory alternative stands out, which leads to a decision-making result that the individual adopts the preferred knowledge (Fadel et al., 2015.).

This research investigates how people evaluate knowledge acquired from online Q&A sites. When people evaluate knowledge, basically they are evaluating the quality of the knowledge. For that reason, a great deal of prior research regarding Information Quality (IQ) was reviewed. For example, the research done by Ge and Helfert (2006) drew a comprehensive big picture of IQ research by summarizing others' relative research work. From that research, IQ was categorized into several groups according to different IQ tasks (Ge & Helfert, 2006). IQ assessment is one of those categories, and it is helpful for answering the research question about knowledge evaluation in this research.

The research question of this study is evaluating the knowledge acquired from online Q&A sites. The ultimate goal of knowledge evaluation is to help individuals make a decision for adopting the most satisfactory option. Hence, the prior research about decision making is useful for understanding this research topic. For instance, Betsch and Glöckner (2010) proposed that decision making process contains two types of processing including intuitive processing and analytic processing. From that research, intuitive processing and analytic processing raised in context of decision making are relatively similar to the processing types from dual process theory. Moreover, intuition is often associated with decision making (Oliver, 2007; Rew, 1988). It is regarded as a valid source and also important component of decision making (Oliver, 2007). For that reason, intuition may also be involved in people's knowledge evaluation on online Q&A sites. Hence, the research regarding intuition in the field of decision making is carefully reviewed.

Dual process theory posits two distinct cognitive systems underlying thinking. Quite a little of contemporary research have evidenced that dual process theory can be applied to contexts of judgement and decision making (Evans, 2003). People's knowledge evaluation on online Q&A sites contains both judgement regarding answers and decision-making for adopting an answer. Consequently, literatures regarding dual process theory should also be reviewed.

Thus, dual process theory and information quality related literatures are reviewed in detail in this section.

2.1 Dual process theory

Dual process theory proposes that information processing can arise through two different ways, which are central routes and peripheral routes (Bhattacharjee & Sanford, 2006). In addition, in most information evaluation circumstances, both routes are involved in filtering process (Chaiken & Maheswaran, 1994). Central route is often used for processing message-related information (Bhattacharjee & Sanford, 2006). Accordingly, central route processing is more explicit and conscious, because it requires more deliberate evaluation based on the information content itself. Therefore, it is usually more time-consuming. Information seekers have to make more cognitive effort before they can make a decision for adopting the high-quality information using central route process

(Fadel et al., 2015.). In contrary to center-route process, peripheral route processes only peripheral cues such as credibility of the knowledge source (Bhattacharjee & Sanford, 2006). It often takes less cognitive effort and time. Thus, peripheral processing is comparatively quicker route (Bhattacharjee & Sanford, 2006; Fadel et al., 2015).

Dual process theory also suggested that the trigger between central and peripheral routes depends on the motivation and expertise of knowledge seekers (Fadel et al., 2015). When knowledge seekers are highly motivated and with high expertise level, they will more likely rely on central cues. That means in order to retrieve more accurate knowledge with better quality, the knowledge seeker needs to analyze the merits of the knowledge on his/her own. If individuals' motivation and expertise are low, peripheral process stands a good chance to be triggered. The reason why peripheral processing is used under this circumstance is that quickly acquiring the knowledge is dominant to the knowledge seeker. In addition to that, the knowledge seeker with low expertise is likely not capable of evaluating knowledge based on its content.

In the context of online Q&A sites, when an individual seeks for the answer to a problem, the individual has to critically analyze the content of the answer if central process is employed. In this process, the individual has to evaluate the answer based on its attributes. In order to do that, the individual has to at least read through the answer content very carefully. Taking Stack Overflow as an example, questions asked there are mainly about programming. Many answers on Stack Overflow contain code example. In this case, the individual has to check whether or not the code example attached to the answer can successfully run to the desired result. However, when the individual encounters peripheral process, the external clues will be taken into consideration for evaluating the answer. In practice, it could be some indicators that show the expertise level of the answerer, which is a common indicator among most online Q&A sites. On programming Q&A site Stack Overflow, there is even indicator showing the amount of vote for an answer. Logged-in users can vote for an answer. Up vote will add one point to the total votes of an answer, whereas down vote will minus one point from the total votes of an answer. Such indicators can reflect the trustiness and usefulness of the answer to some extent, even though they may not reflect the real quality of the answer (Fadel, et al., 2015).

Some prior research asserted that central process and peripheral process may occur alone, or can co-occur in certain contexts (Chen, Duckworth & Chaiken, 1999; Fadel et al., 2015). Individuals can freely shift between central and peripheral routes depending on the ongoing evaluation task (Fadel et al., 2015). The finding was believed to be adaptable in the context of Stack Overflow as well. It is because of two reasons. On the one hand, in Stack Overflow, there are different indicators displayed next to an answer. Those indicators consist of votes of the answer, author reputation score, accepted answer mark etc. Both of votes of answer and author reputation score are numeric value that can be quickly checked without taking much cognitive effort. Accepted answer mark is a green tick that is shown at the left side of the answer, just below the votes of answer. Accepted answer mark can be easily noticed as well. Based on that, those indicators are actually peripheral cues that site users can employ when evaluating the knowledge. Since there are several peripheral cues around the answer, it is likely that site users can be aware of that. When site users take such peripheral cues into the knowledge evaluation process, the peripheral route is adopted. On the other hand, the purpose of using Q&A sites is to find a proper solution to solve the real-world problem. That means site users have to read through the content of the selected answer in order to understand and implement the solution mentioned in that answer. At this point, central route is employed by site users.

2.1.1 Central route: information quality theory

The research question is about how people evaluate knowledge retrieved from online Q&A sites. When people take the central route to evaluate knowledge, they are actually evaluating the quality of the content itself. For this reason, some prior research regarding information quality is discussed below.

Ge and Helfert (2006) reviewed plenty of previous literatures regarding information quality. They discovered that information quality research results have been applied to many varied industries, among which information system and decision making are mostly cited. This finding double confirms that information quality research can provide some hints for this study.

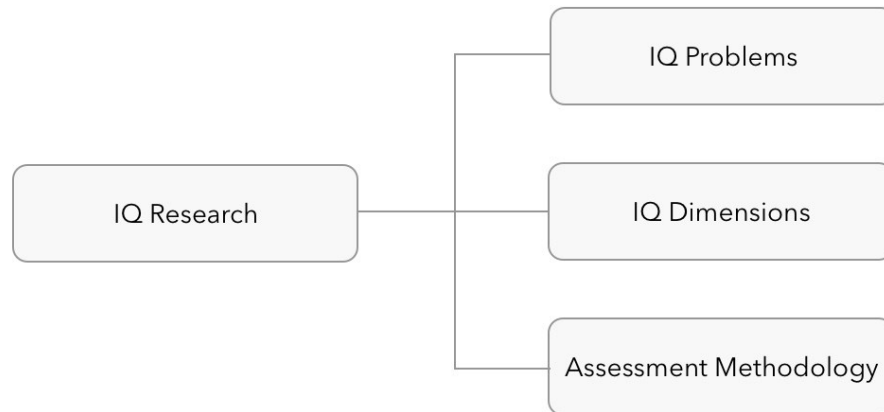
On the one hand, online Q&A site is a type of small information system. Zwass (2017) defined information system as a collection of integrated components that gathering, storing and processing data, meanwhile, supplying knowledge and information to its users. While on online Q&A sites, users create, communicate the information over the website, which is the input of the system. Website stores all the information on the server; this is the storage of system. Website organizes all the information properly, and users can also search from website database for previous questions and answers; this is the processing and control of information. All information on the website is retrievable for users and search result is shown to users as well, which is the output of the system. Based on all the factors mentioned above, online Q&A can be regarded as an information system accordingly. On the other hand, the research question is evaluating the knowledge acquired from online Q&A sites. The reason why users from online Q&A sites evaluate knowledge is to obtain one or several answers that satisfy their needs. Obtaining one or more answers is a decision-making matter. Based on the two aspects illustrated above, we believe that information quality research would light up the road for this research.

Ge and Helfert (2006) structured information quality (IQ) into three dominant categories: IQ assessment, IQ management, and Contextual IQ. IQ assessment can be defined as the process of evaluating information quality based on different IQ dimensions. The goal of IQ assessment is to understand how content from information system is assessed. IQ management contains management for information, knowledge and quality. The purpose of IQ management is to improve the information and knowledge quality. Contextual IQ research inspects the effect of information quality under different organizational contexts such as World Wide Web, information manufacture system and finance (Ge & Helfert, 2006.).

IQ assessment can be expanded into three more detailed sub-categories, which includes: IQ problems, IQ dimension, and IQ assessment methodology (see Fig. 1). IQ problems research investigates various problems of information, such as bias and out-of-date issues. IQ dimension defines the factors that affect information quality. IQ assessment methodologies can be defined as integrated approaches designed for assessing information quality (Ge & Helfert, 2006.).

The current research scrutinizes how people evaluate knowledge retrieved from online Q&A sites. This involves an assessment process regarding information quality. This research focuses on investigating the process of knowledge evaluation, in which IQ problems are not dominant in this research. For that reason, problems of IQ are not much relevant to this research. Therefore, it will not be explained into details. IQ dimension which defines multiple dimensions that can be used for assessing information quality, and

IQ assessment methodology which defines the means to assess information quality are extremely relevant to this research and thus they are discussed in the following.



Conceptualized based on Ge & Helfert (2006)

Figure 1. Information quality research categories

Ge and Helfert (2006) suggest that information quality is a multi-dimensional concept; each dimension can affect IQ assessment to some levels. Zhu, Bernhard, and Gurevych (2009) proposed a dimension-based quality model which comprises thirteen dimensions including: informativeness, politeness, completeness, readability, relevance, conciseness, truthfulness, level of detail, originality, objectivity, novelty, usefulness and expertise. Brief definition for each of dimensions is presented in Table 1. Shah et al. (2010) used these dimensions as the evaluation criteria in their research after elaborately going through several correlated prior literatures. The reason why they adopt Zhu et al. (2009)'s thirteen criteria is that it is the most fully developed set, which indicates that the dimensions summarized by Zhu et al. (2009) is trustable.

From a different point of view, IQ assessment can also be categorized into objective assessment and subjective assessment based on its target (Ge & Helfert, 2006; Pipino, Lee, & Wang, 2002). Objective assessment focuses on data set, in which the natural properties of the data are measured and compared with the optimal value. For that reason, it is comparatively easy to run the objective assessment automatically with help of software. However, subjective assessment targets at data consumer, and is commonly used to measure to what extent the information satisfies the data consumer for its usage. Thus, it is improper to define a rule for running an automatically measurement because of the subjective attribute. As a result, in this situation user survey is commonly used instead. Each individual may have different standards from each other regarding information quality. Consequently, even if several people assess the completely same information, because of subjectiveness, it is still possible to generate different results.

In the context of online Q&A sites, the behavior of knowledge evaluation conducted by users is mostly subjective for two reasons. First of all, it is the user who assesses the quality of knowledge. No matter what standards the user uses for assessment, that is all from the user's subjective consciousness. Secondly, on many online Q&A sites, there are various attributes that indicating the quality of the knowledge; for examples, the votes for an answer shows its usefulness to site users. The more votes an answer receives, the more likely it is accepted by other users in the community. Yet, it is still the user who votes for

an answer. The vote itself is a result of several added subjective decisions, thus it cannot be used as an objective rule for measurement.

Table 1. Information quality dimensions and definitions

Conceptualized based on Ge & Helfert (2006)

Dimension	Definition
Informativeness	Whether the amount of information included in the answer is enough to answer the question
Relevance	How suitable is the answer to the context of the question
Truthfulness	The level of trustworthiness of the answer
Readability	How easy the answer can be read
Conciseness	How tidy is the appearance of the answer
Usefulness	How helpful is the answer to address the question
Level of detail	How many details the answer contains
Expertise	The possibility that the answer-writer is an expert
Originality	How much of the answer content is created by the answer-writer himself/herself
Completeness	How thorough the answer can reply to the question
Objectivity	How neutral the answer is
Politeness	The extent that the answer reflects the answer-writer's care regarding others' feelings
Novelty	How creative the answer is

2.1.2 Peripheral route: intuitive processing

When using "ajax form submit" as search keyword on Stack Overflow, it gave a result of 64,123 candidate threads. This result contained a large set of alternatives, which was almost impossible for users to go through all of the candidate threads. Thus, one more word "jQuery (jQuery is a JavaScript library)" was added to search keyword to possibly reduce the amount of candidate threads in search results. When searching with reorganised keyword "jQuery ajax form submit", there were still 25,844 alternatives (data was retrieved on September 12, 2017) shown in search results. It was still impossible for users to go through each piece of search results and systematically evaluate its quality. However, users can evaluate the knowledge based on some peripheral clues. As introduced in dual process theory, peripheral process is typically less time-consuming and less effort-costing. Thus, conducting peripheral process might significantly reduce the time of knowledge evaluation on online Q&A sites, and also take less cognitive effort from users, therefore make the knowledge evaluation processing easier for users.

Similar to peripheral processing, intuitive thinking also has the benefit of taking less cognitive effort while still processing information quickly (Betsch & Glöckner, 2010). Betsch and Glöckner (2010) proposed three main characteristics of intuitive processing

in their research. The first remarkable characteristic is autonomy, which means intuitive processing can be triggered automatically without intentional control. Secondly, intuitive processing is much less constrained by the cognitive capacity and amount of information. That means even if more information is awaiting to be processed, the time it costs may not be more, but may be even less if the coherence of entire pattern is increased. The last main characteristic is that intuitive processing can process multiple pieces of information in parallel and in a relatively narrow time frame too (Betsch & Glöckner, 2010.).

The research done by Honsel, Herbold, and Grabowski (2015) indicated that users have considerably good intuition towards postings on Stack Overflow. The research investigated 9 myths about getting an answer for the question and voting behavior. The 9 myths come from a result of few developers' brainstorming, therefore those 9 myths show how those users see the answers on Stack Overflow to some extent. Regarding the voting behavior, developers expect it might be affected by length of answer and whether there is source code enclosed. The result shows developers were actually right. It would be interesting to see whether those factors that affect voting behavior would also affect individuals' evaluation towards answers.

Analytic process is another way of thinking that contrasts intuitive process. Analytic process is much alike central route thinking from dual process theory. Both analytic and intuitive processes work together in information evaluation and decision-making process. Intuitive process is dominant when integrating information and output the decision are the main tasks. Whereas analytic process is mainly responsible for input information such as information search and change (Betsch & Glöckner, 2010). The research investigates individuals' knowledge evaluation process on online Q&A sites, which involves knowledge search and the decision-making process to adopt the knowledge. Therefore, both analytic and intuitive processes might be included.

3. Methodology

According to Hoepfl (1997), qualitative research normally observes perception of participants regarding their experiences in certain context. Patton (1990, p. 55) also stated that qualitative research is typically about investigating and elaborating the patterns of things in its natural settings, without changing the data. Comparing with quantitative research, which is about figuring out “what, where, when, and how many” in general; qualitative research is about investigating “why” phenomenon (Brenner, Brown, & Canter, 1985, p. 116).

This research aims at understanding the process by which people evaluate knowledge on online Q&A sites, which is about learning the users’ experiences. Subjective opinions from participants make up the main body of the research data. All of those opinions should be kept in its original form. That is because in order to answer the research question, the research requires a summary of all factors that participants consider during knowledge evaluation on online Q&A sites. Even though the research question is not directly asking a “why” question, it is still for understanding a phenomenon about people’s behavior in knowledge evaluation process. Hence, a qualitative research should be conducted.

This research investigates mainly the subjective opinions from participants. For example, participants were asked about what factors they thought that would affect their knowledge evaluation process. For such open question, the participant could give any answer that might contain one or a lot more factors. Since each individual has different levels of organizing oral expression, there might be a chance that some expression was not clear enough to be understood by the researcher. Interview is helpful for avoiding such obstacle. That is because if the researcher notices that one answer from the participant is not fully understandable, then the researcher can immediately raise another question to investigate more into the answer (Brenner et al., 1985, p. 3).

From the interview, interviewees were asked about what factors that would affect their knowledge evaluation process as part of prepared questions. Since there were no predetermined factors, hence the factors were defined totally based on interviewees’ own experiences. It is possible that the researcher might think that more interpretation is needed regarding some factors. In such situation, the interviewer should add new questions for the unclear specifications. Consequently, semi-structured interview should be implemented. Because semi-structured interview is defined as the interview that has an incomplete script, where improvisation might occur during the interview (Myers, & Newman, 2007).

3.1 Research site

In this research, Stack Overflow is used as the research site. This research is made to investigate how do people evaluate knowledge acquired from the online Q&A sites? There are three reasons behind this decision of taking Stack Overflow as research site. First of all, it accords with the research object that it is one of those online Q&A sites. Secondly, it is the largest online community for developers. Up to January 22nd 2018 by 8:24pm (UTC+2), there were 15,229,369 questions asked on Stack Overflow, with 4,374,439 questions had at least one answer. Thus, it could provide large enough dataset for the research. With larger dataset, the result can reflect more general implications. Thirdly, Stack Overflow is one of the most popular online Q&A sites among developers with more than 50,000,000 professionals visit monthly (<https://stackoverflow.com/>).

Therefore, we believe that after studying this site, the findings generated through this research would benefit online Q&A sites to the direction that how the site can make user interaction smoother and to better serve the users, so that this research would be helpful for millions of developers who are using the site.

Stack Overflow is used as a platform for site users to solve programming problems. Site users could rise a question, and others from the same community may add an answer under the question. For many threads, there is more than one answer. Sometimes a question could even receive over twenty answers.

On Stack Overflow, some very old threads still kept receiving new answers long time after the question was asked. From this point of view, Stack Overflow can be used as a storage of documentations for many common programming issues (Gantayat, Dhoolia, Padhye, Mani, & Sinha, 2015). Even if someone started learning a long-existing programming language, the individual can always find some common questions from old threads, and check the answer history of that question, whereas some new answers might correct something that is no longer valid from old answers.

3.2 Data collection

Myers and Newman (2007) suggested that researchers should try to avoid the elite bias when using qualitative interview. Elite bias occurs when data is collected mainly from high status informants rather than the low status informants (Heiskanen, & Newman, 1997). It is recommended that researchers should try to contain variety in the data samples (Myers, & Newman, 2007). Therefore, in order to abate the elite bias of collected data, participants for interview are randomly selected without asking for any prerequisite. However, since the interview is designed for understanding user experiences over Stack Overflow, there should be a constraint condition when choosing participants. The participants should have prior experience of using Stack Overflow.

The way researcher looked for potential interviewees was through directly asking whether or not the individual would like to participate in an interview that was for the research about how people evaluate knowledge over online Q&A sites. After the individual agreed to be involved in the interview, an interview method (either online interview or face-to-face interview) was set. If the participant prefers online interview, then an online interview tool was agreed in advance. If the participant prefers face-to-face interview, then a time and place were set for the interview in advance.

In this research, sixty-seven people (57 males, 10 females) were interviewed between September 2017 and January 2018. Because of geographical and time constraints, most interviews were done through online chatting tools. Regarding the demographical information, twenty-three participants are in age 18 and 24 (see Fig. 2). Twenty-nine participants are in age 25 and 29. Thirteen people belong to age range 30-34. Only two people are in age range 35-39. Geographical locations of participants were distributed among: Finland (49), China (16), USA (1), and Australia (1).

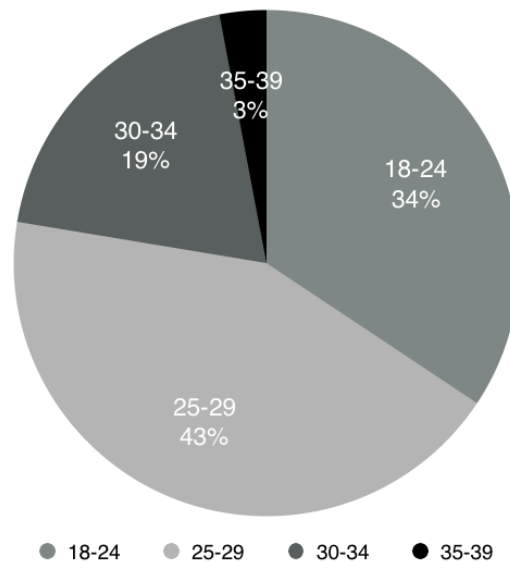


Figure 2. Age distribution of participants

Participants in this research also have different educational backgrounds: Bachelor's Degree (34), Master's Degree (21), Doctor's Degree (1), high school (9) and vocational school (2). They are university students, Ph.D. candidates, university staff from IT research unit, employees from IT or software companies. In the interview, the participant's working experience in field of Information Technology was also investigated. Nine participants have no previous working experience at all; forty-nine participants had no more than five years working experience; seven participants had more than five years but no more than ten years working experience; only two participants had more than ten years working experience.

Most participants use Stack Overflow to find answers to programming questions. However, there were a few participants also reading discussion on Stack Overflow for browsing the latest technologies and other useful knowledge. Moreover, one participant used Stack Overflow for job seeking as well. Majority participants (54) have had experience of using Stack Overflow for no more than five years, counting from the first time use of Stack Overflow to present. Only thirteen participants were comparatively more sophisticated users who have had more than five years experiences. Besides, twenty-six participants had actual account on Stack Overflow while others were just using it anonymously.

3.3 Interviews

The interview used in this research is designed as semi-structured interview. The interview questions are open-ended. One advantage of adopting open-ended questions is that interviewees are allowed to freely and fully express their own experiences in details, which is helpful for enriching the research data. Another advantage is that using open-ended questions can weaken the bias from researcher, and this effect would be even stronger when there is large number of interviewees (Turner, 2010.).

An interview was conducted with each participant. It was either done through some online chatting software or through face-to-face interview. Most of the interviews were done via online chatting software. For a few participants who located at the same city as the researcher, a suitable time and place was set up for conducting the face-to-face interview.

(However, many participants who were living in the same city as the researcher still preferred to use online interview).

There are two reasons why most interviews were conducted through online chatting tools. Firstly, online chatting tools have advantage of easily reaching people who are difficult to access (Opdenakker, 2006). For instance, people locate at other countries are not able to participant in face-to-face interview. Secondly, since the interviewer and interviewee could not see each other, therefore some social cues are hidden. Interviewees who are bashful in face-to-face contexts may feel safer and more confident to share their experiences (Rheingold, 1994, pp. 23–24.).

All the interviews that were conducted through online chatting tools were done synchronously. According to James and Busher (2012), synchronous interview has advantage of real-time interaction between interviewer and interviewee. That is similar to the traditional face-to-face interview. Because of the nature of real-time, the interviewee can be more involved in the interview. Moreover, such immediate conversation can also help interview participants get to know each other better, consequently weaken the psychological distance between participants (James & Busher, 2012.).

The online chatting tools used for interviews consist of Facebook, WeChat, Mattermost, and QQ. Facebook is world famous social networking website, and it also has mobile application. Not as well-known as Facebook by the world, WeChat is a social networking mobile application that people use for free text message and calling, it is the most popular social networking application in China (Boyd, 2017; WeChat, 2014). In this research, most of the interviews for Chinese participants were done via WeChat. Mattermost is an open-source communication tool, which can be installed both on computer and mobile (Mattermost., 2018). QQ is a messenger software developed by Tencent, and it also has mobile application (Imqq.com, 2018).

Comparing with taking notes while doing interview, keeping recordings of interview can benefit in remaining the faith of the retrieved data. Furthermore, it can also help the researcher from deconcentrating during the interview (Hoepfl, 1997.). Because of the benefits that recording interview can bring, all face-to-face interviews done for this research were recorded. During the interview, a voice record was taken after getting approval from interviewees, then the researcher wrote down the answers into recording document afterwards. If the interview was done over online chatting software, after interview, the researcher pasted all the original answers into recording document.

The main question for the interview is: *How do you evaluate the knowledge and choose the best one? In other words, what are the factors that affect your evaluation and decision on adopting certain information or knowledge?* In addition, other questions such as the demographical information of participants (age, geographical location, education, work experience and occupation) and how often they use Stack Overflow were also asked. There was no constraint for participants answering these questions. And the participants can freely use words to answer questions. All questions could be regarded as neutral since there was no judging words used. During interview, researcher asked only one question at a time. There was no “why” question asked in the interview. As noted by McNamara (2009), “why” questions might bring a cause-effect that may not exist. It might also make participants feel cautious (McNamara, 2009).

The interview content covered not only some key questions related with using Stack Overflow and knowledge evaluation process there, but also some simple questions that

investigating participant's background information, such as educational background, related working experience, and frequency of utilization etc. Those background related questions are designed for assisting the researcher to find some other factors that might affect knowledge evaluation on online Q&A sites, to see if there is any correlation between knowledge evaluation process and those factors. Those questions are open questions, since it is not known beforehand what factors the interviewee could have in mind. Questions for investigating the interviewee's background information are mostly closed questions, and those are designed to be short. This is because too long interview might bring discomfort to the interviewee and also such demographical information is not dominant in this research.

After interview phase was finished, all data retrieved from interview were summarized and documented. The dominant data were the factors that participants used for evaluating answers from online Q&A sites. Thus, these data were not only summarized, but also categorized accordingly, from which the researcher tried to discover if there were any common key factors that people employed, and how differently people evaluated the knowledge.

4. Data analysis and results

Interview recordings were coded based on interviewees' answers. Such typed transcripts are used quite often in interview research (Mack, Woodson, MacQueen, Guest, & Namey, 2005, pp.30). In this research, each interview recording was also typed into individual documentation stored on the computer. Those documents were the main materials for data analysis.

Data analysis phase started after data collection phase. In this phase, data that were gathered from interviews should be illuminated and interpreted into categories or groups of information (Turner, 2010.). This process of data interpretation is content analysis. This data analyzing method is commonly used in qualitative research because of the subjective nature of the data collected for research. The purpose of content analysis is to make sense of the raw materials, so that ultimately both of the surfaced and implicit meaning of the research data can be understood (Brenner et al., 1985, pp. 117-118.).

Elo and Kyngäs (2008) asserted that the outcome of content analysis are categories which could illustrate the research context. They also introduced that content analysis could be either inductive or deductive, depending on the purpose of research. Inductive content analysis is used for describing a generic view of the phenomenon. This generic view is produced through analyzing and combining all the detailed and specific research data. For that reason, data is moved from specific to generic. In contrast, deductive content analysis is mainly used for theory testing. Therefore, deductive content analysis commonly has a based theory or prior knowledge. Thus, the data analyzed in deductive content analysis goes from generic to specific (Elo & Kyngäs, 2008).

This research addresses itself to having an overview of different factors that affect people's knowledge evaluation process on online Q&A sites. Factors that each participant states are the very specific and detailed information. However, when taking all participants as a collective, these data may become generic. It can be seen that the data categorization develops from specific to general. Thus, the inductive content analysis is more proper for this research.

There are three steps that Elo and Kyngäs (2008) have suggested for inductive content analysis, which includes open coding, categorization, and abstraction. Firstly, the researcher should go through all the materials again. The purpose of this step is for the researcher to fully understand each answer. Secondly, the researcher should identify categories that are emerged from the interview answers. This could help researcher to better illustrate and understand the research phenomenon. Thirdly, categories generated from the second step should be evaluated. Categories with similar characteristics should be combined into one generic category. Some generic categories might also be combined into one main category (Elo & Kyngäs, 2008.).

Following the instruction given by Elo and Kyngäs, the interview materials were coded after interview phase in this research. The key question for the interview was: *How do you evaluate the knowledge and choose the best one? In other words, what are the factors that affect your evaluation and decision on adopting certain information or knowledge?* In order to analyze the answers for this question, all factors emerged from participants' answers were coded and grouped into categories. The categorization process occurs at the same time while inserting data into the summary document. A few examples of categorization will be illustrated in the following.

For example, one participant said: “I read those answers and check how many points the answer is received”. The other participant said: “I check how many votes the answer already has”. Basically, both participants are talking about votes of the answer, therefore these two answers can be both placed under category “answer votes”.

Another example: one participant stated: “usually the good answer has the characteristic of long answer with very detailed explanation”. While another one stated oppositely: “I will take the answer which is shorter answer”. Even though one identifies long answer as good answer, the other one prefers shorter answer, both are linked to lengthiness property of the answer. Thus, these two answers were categorized into the same category: lengthiness.

Each participant’s responses were examined. If common characteristics were found among a few answers, then those answers were combined into same category. If new characteristic is found from an answer which has not emerged before, then a new category is added. This categorization process continued until all the existing categories were saturated and no more new category emerges.

After the first categorization phase, twenty-two generic categories were defined (see Fig. 3). After carefully examined those twenty-two generic categories, the researcher found there were still similar characteristics among certain categories. Therefore, another categorization process was conducted for generating more refined categories. The process will be introduced in the following.

Categories “code example”, “concise and clean code”, and “length of code” are all about code example, therefore, those can be merged with category “living example”. Code example is specific to programming related online Q&A sites (e.g. Stack Overflow), thus it may not apply to other types of online Q&A sites which is not relevant to programming. Therefore, the factor is entitled as living example in order to be more generic.

“Votes”, “accepted answer”, “votes of comment”, and “amount of comments” are different types of indicators of an answer, thus a main category “answer indicators” can be formed to cover all these similar categories. “Confirmation to own knowledge” and “meet own expectation” are both based on one’s former experiences, hence these two can be combined to category “empirism”.

Feasibility is defined as the possibility that a task can be done, or a goal can be achieved, or something is reasonable in Cambridge Dictionary (Dictionary, f., 2018). Participants explained they concerned whether the answer contains similarity to their own contexts, because this will ensure the solution mentioned in the answer can be applied to their circumstances. “Solution is tested to be working” ensures the desired result can be achieved. “Complexity to apply the solution” is talking about to what degree the solution can be applied in practice, which is about feasibility apparently. While participants consider “comments from others”, because they tried to find clues from others’ comments to prove the feasibility of the answer. In conclusion, these four categories can be placed under main category “feasibility”.

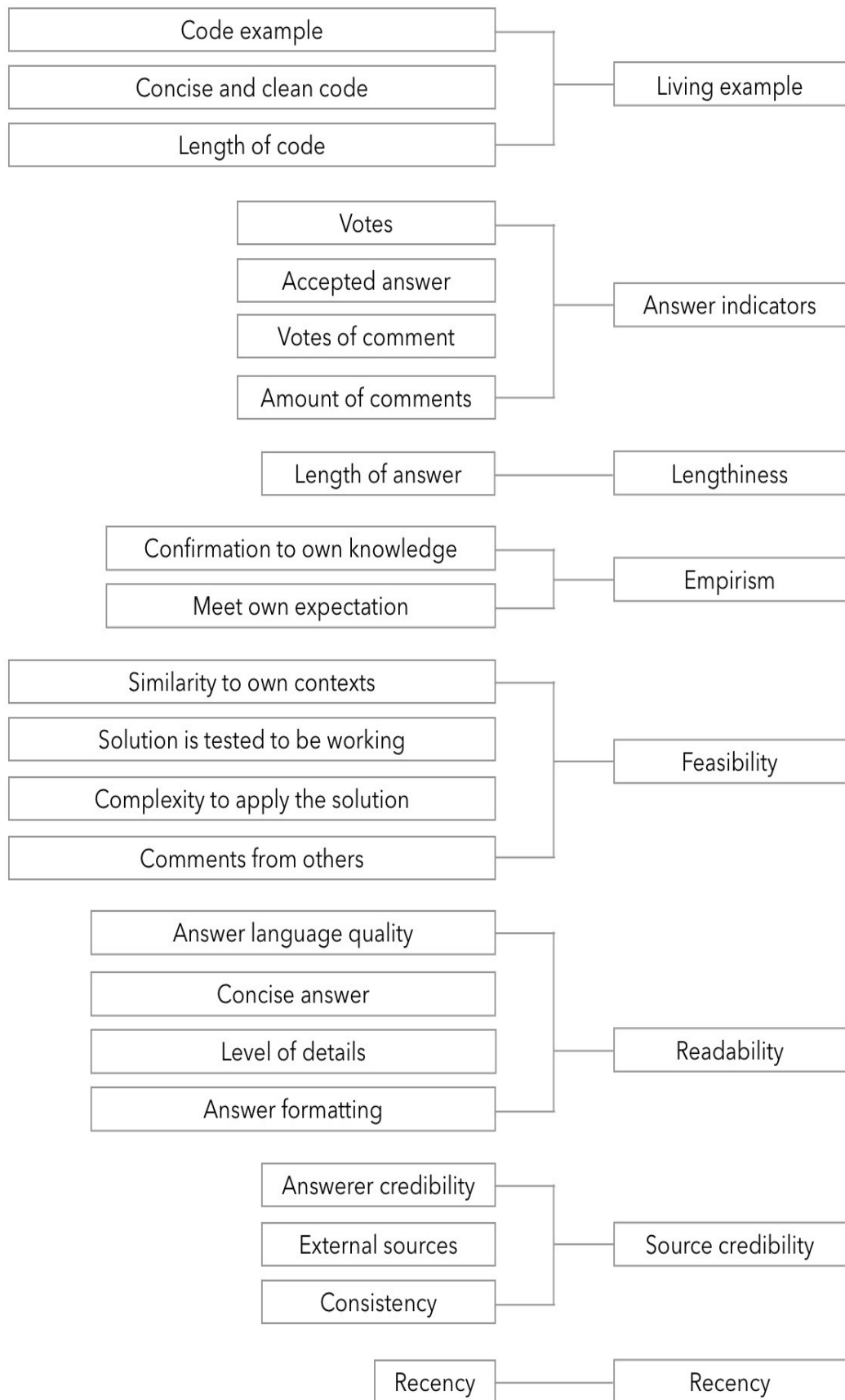


Figure 3. Main factors formation

“Answer linguistic quality” mainly refers to the linguistic quality, for instance grammar correctness and word accuracy. “Concise answer” requires the answer writer to have ability of expressing their ideas in a clear and accurate way. “Level of detail” depends on writer’s ability of literal content organizing, which shows writer’s writing ability. “Answer formatting” would affect the quality of readability as well. If text is not formatted well, then it will be difficult for reading, which will decrease the quality of the answer to readers. All of these categories mentioned above is relevant to answer writer’s writing quality. However, the answer writer’s writing quality would actually affect the readability of an answer. For that reason, those categories were combined into a main category “readability”.

“Answerer credibility” is employed by some participants when evaluating the answer. They illustrated that if the answerer has good credibility, then the answer given by the person seems to be more trustable. Therefore, the answerer credibility is the one factor assists for judging the source credibility of answer. Participants take “external sources” into account as well, because they hope to find some clues from those external sources which proves the correctness of the answer. Thus, this is also about answer’s source credibility. Moreover, some participants explained that they prefer to accept the answer, if “several people give the same answer”, because the consistency among different answers made them feel the knowledge included in the answer is more reliable. As mentioned above, those three categories are all about the trustiness of an answer, therefore these can be merged to one main category “source credibility”.

4.1 Main factors that affect individuals’ knowledge evaluation process

After second categorization process, “lengthiness” and “recency” were found not having any common characteristics with any other categories, therefore, these two categories were remained without merging with any other categories. In the second categorization phase, there were totally eight categories were formulated (see Fig. 3), these are: living example, answer indicators, lengthiness, empirism, feasibility, readability, source credibility, and recency. Each factor is explained along with some representative examples in the following.

Living example

Since most questions asked on Stack Overflow is about programming, therefore, it stands a good chance that coding example is involved in the answer. Actually, many answers on Stack Overflow contain a code snippet. Some participants expressed that they prefer the answer with code snippet since it is usually more illustrative. One participant illustrated that “I like answers with snippet, with a live demo”. In addition, some of participants will also check some attributes of the code example, which includes the length of code, formatting of the code, and whether the code is concise enough. For example, one participant explained: “I prefer the answer with clean code if I’m looking for a code example”. They also explained that those attributes can reflect the quality of a code example to some levels.

Answer indicators

Answer indicators refer to the index that shows others’ feedback regarding the answer. Answer indicator is the result of a collective evaluation or feedback from users in the community. In Stack Overflow, those indicators consist of: votes of answer, accepted answer, votes of useful comment, and amount of comments. The reason why participants

look at those indicators is that those can reflect the quality of the answer to some extent. Like one participant illustrated: “I think the answer that has more votes is more reliable”. Some participants also thought “the answer which has more comments seems more valuable”.

Lengthiness

Several participants consider length of an answer as one evaluation criteria. Some of them prefer longer answer. For instance, one participant explained like this: “usually the good answer has the characters of long answer with very detailed explanation, and offer practical solution”. In contrary, some others like the shorter answer because shorter answer takes less time to read through. One participant also introduced that even if the short answer is good, but it should not be too short to explain the solution: “Questions doesn't need long answer so I skip long answers, unless they have the short answer in beginning and only then I may read whole text. Same goes with too short answer. One sentence usually doesn't answer to my problems.”

Empirism

Some participants explained that they will choose the answer which confirmed the knowledge of their own, or met their expectation. That is more rely on tuition without any implemental testing or prove. One participant said he will adopt the answer if “it answers the question the way I feel is correct”. Another one suggested that “I try to search something similar to what I expect”. It can be seen that individuals judge the answer based on their prior knowledge and experiences even if there is no clear evidence to support this judgement.

Feasibility

Many participants stated that the most important thing for them is that the solution proposed in answer is actually working. The answer might be tested by themselves or someone else. One participant explained “the best way to evaluate the answer is to implement it to see whether it's correct or not”. Another one said: “I consider answers that the developer said something like: ‘I tried and worked for me’ or something”. Several participants also explained that they can find prove of feasibility from comments left by other users to the answer. As one participant illustrated “there're usually more clues from those comments”. The useful information sometimes is the evidence that proves the solution suggested is tested to be working. Yet sometimes it could even be the evidence indicating a failure solution.

Readability

Readability is important for site users when evaluating an answer. Since readers need to understand the content of the answer first, before they can adopt the solution mentioned in the answer. Writing quality can affect the understandability of the answer to a great extent (Powell & Hebert, 2016). One participant explained: “Answer description text should be clear and easy to understand”. According to participants, there were several aspects of writing ability: grammar accuracy, conciseness, level of detail, and formatting of text.

Source credibility

Credibility of both of answer author and answer content are in consideration of some participants. The reputation score is shown under answer author's name on Stack

Overflow. Some participants mentioned that they usually check that indicator of the answer author. They thought that the higher score the author has, the more likely the answer given by that user is trustable. In addition to that, source credibility of answer content is also an important factor to some participants. External sources included in the answer and also the consistency to answers from others can increase the credibility of the answer. One participant suggested that: “if the answerer has included links to sites with more information, or to back up their answer makes me feel like the answerer knows what they are talking about”.

Recency

Twelve participants introduced that they would pay attention to when the answer was given. In Stack Overflow, the date and time of an answer is displayed beneath the answer main content. Date and time of the comment is also shown close to the comment itself. As explained by some participants, recency is particularly important for programming related answers. Because the programming language, software, framework, and plug-ins may have many versions as time goes on. The answer seeker might be looking at an old answer from two years ago. The answer was working for an older version, but the solution is not working any more for the latest version. As one participant stated: “I usually look at the answers posted recently, sometimes the old answer would not work in the latest version”.

4.2 Popularity of main factors that affect knowledge evaluation process

A group of main factors appeared through data analysis. It was extremely helpful for answering the research question about how people evaluate knowledge retrieved from online Q&A sites. Moreover, the researcher thought that it would be ideal to have the statistics that indicate the popularity of each main factor. Because it would be useful for providing a different perspective regarding the individual’s evaluation process. For that reason, how many people have contributed to each factor was counted. One individual’s contribution added one point to that factor.

However, a problem emerged when counting for the points of each factor. When creating main factors, one or several original factors were merged. In that case, there was possibility that one individual might contributed to several original factors under one main factor. Take this example, an individual has stated that both of code example and concise code were in his consideration in knowledge evaluation process. That added two points to the main factor living example, even though the two points were actually from the same participant. In order to avoid such overlapping points, the points of each main factor were rechecked. After rechecking, the correct number of points for each main factor were calculated (see Fig. 4).

What can be seen from the result is that answer indicators (53 points) and feasibility (52 points) are the two most dominant factors. The total sample number in this research is sixty-seven. Over 79% of the participants have referred to answer indicators. Nearly 78% of the participants looked into feasibility as well. Living example (28 points) and empirism (24 points) were the second most dominant factors. Then readability (16 points) followed after. Recency (12 points), source credibility (11 points), and lengthiness (10 points) were about the same level of popularity. From all the main factors, lengthiness was the least dominant factor with only ten supporting references.

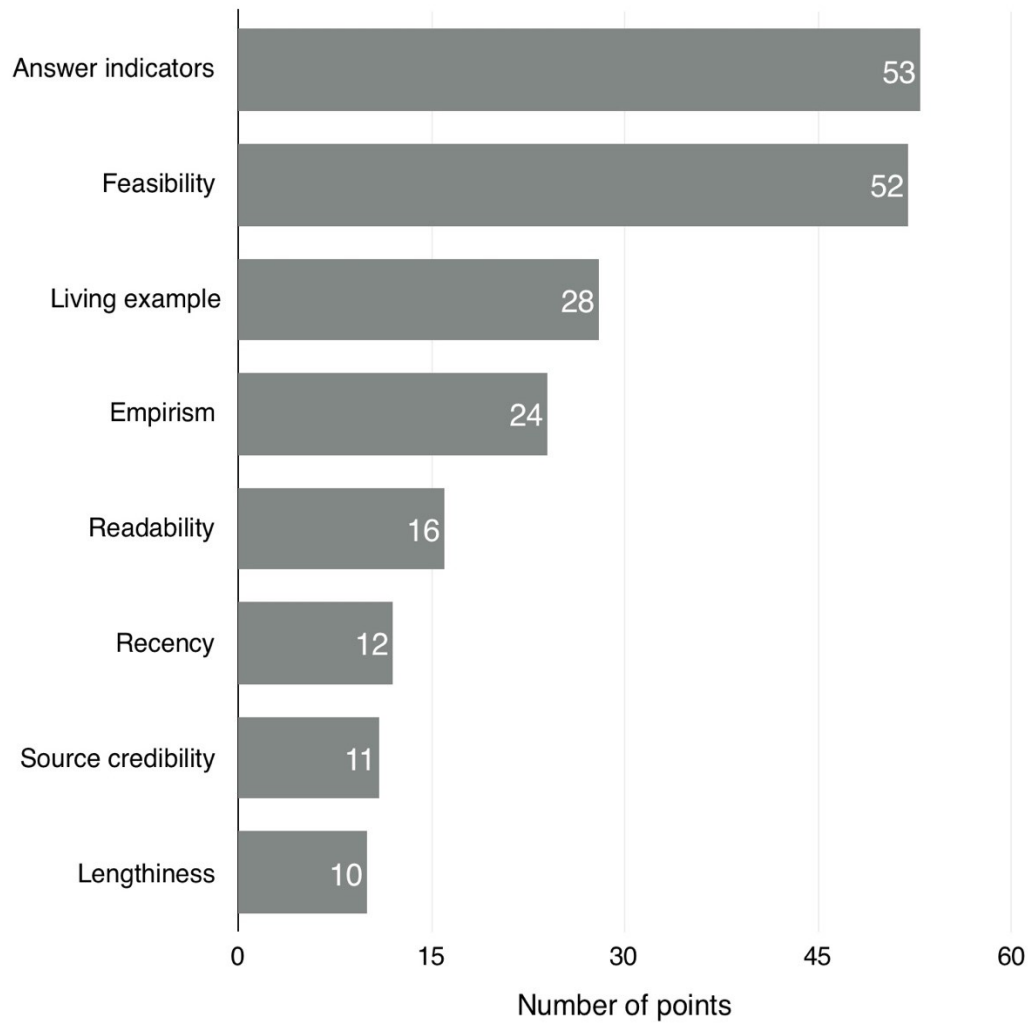


Figure 4. Popularity of main factors

5. Interpretation Based on Theories

In what follows, the factors identified in this research are interpreted based on dual process theory and information quality literatures. The findings confirmed to those theories to some extent.

5.1 Interpretation based on dual process theory

The principal objective of this research is to answer the question of how people evaluate knowledge retrieved from online Q&A sites. Based on the results of this research, in total, eight main factors are involved in individuals' knowledge evaluation process. Some of those factors share similar property. For instance, answer indicators, lengthiness, readability, source credibility, and recency can be described as heuristic cues of an answer (Ratneshwar & Chaiken, 1991). Those factors are relatively easy to employ irrespective of the answer content, because they take less cognitive effort and is time-saving. For example, the individual only needs to glance at those numeric indicators to retrieve information that indicates the quality of an answer to some levels without reading through the answer content. Hence, those factors can trigger the peripheral-route process. However, other factors including living example and empirism are comparatively time-consuming factors that individuals usually have to spend more time and proceed with critical thinking. Therefore, the process of evaluating those factors is a central-route process according to the dual process theory.

This result demonstrates that evaluating knowledge from Stack Overflow involved a dual-process which includes both peripheral and central routes. Main factors and its corresponding process routes are summarized in Fig. 5. Each factor and its relationship with dual process is explained in the following.

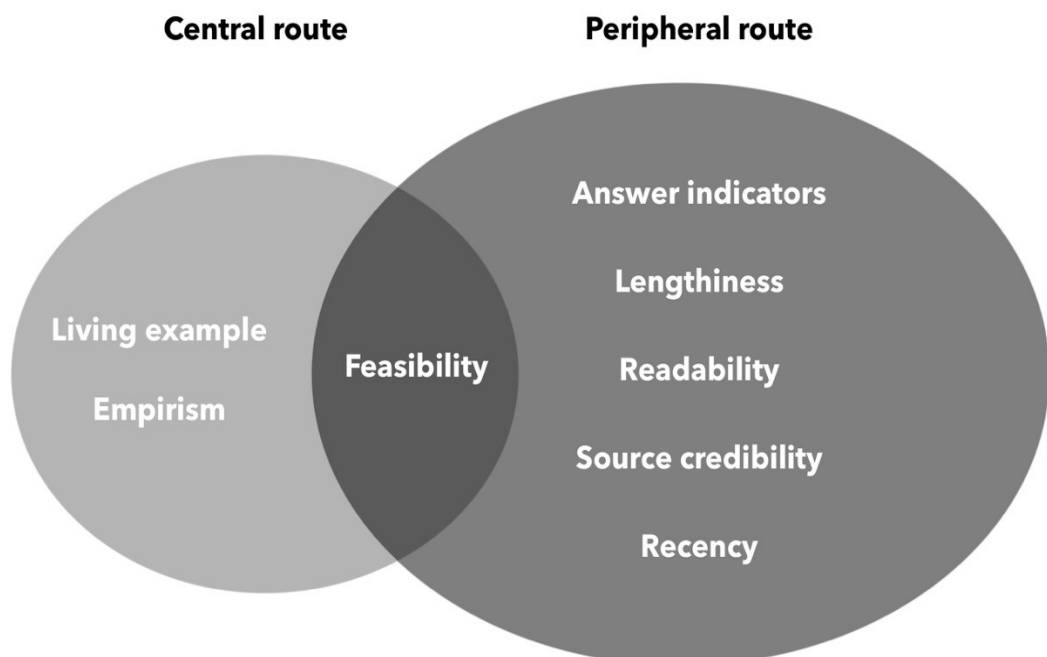


Figure 5. Main factors reflection on dual process

Living example

Code example was mentioned by many participants in interviews from this research. We believe, the reason behind this result is that the domain of Stack Overflow is a programming Q&A. A majority of questions asked there are about programming, where coding is nearly an inseparable part. Code example attached to an answer can make the solution more visualized when the context is to solve the bug from a section of code. Code example contains not only the code, but also rich information. Hence individuals could read through the code example to know what function it executes, thereby evaluating or even testing out if the code could help solve the problem. Good code example is a very efficient tool for developers to solve their programming problems (Nasehi et al., 2012).

Living example is a part of the answer content itself. Individuals have to read through the example content to check its validity. One of the distinctions between central and peripheral route is that the former requires heavy working memory (Bago & De Neys, 2017). Checking code example requires individuals to load knowledge from their memories, therefore it triggers central route.

When participants explain code example related topics, some of them also proposed that concise code was important. One participant also explained that he preferred answer with short code example. Nasehi et al. (2012) introduced that concise code should have attributes of less complexity, and with short code example. Furthermore, concise code is one attribute that contributing to recognized answer (Nasehi et al., 2012). That illustrated code example can indeed affect the quality of an answer, and concise code with short code example reflecting better quality. Therefore, individuals can take living example as heuristic cue to evaluate the quality of the answer, and it might lead to good result.

Answer indicators

The vote number of an answer shows how useful the answer is. Votes of the comment also shows how useful that comment is. The accepted answer is chosen by the question asker when s/he thought one answer is the best among all others. The selection of best answer shows the satisfaction from the asker, which means the best answer should meet the asker's information need (Liu, Bian, & Agichtein, 2008). In terms of online Q&A sites, it means the answer is able to address the problem. According to a research done by Gantayat et al. (2015) which investigated the synergy between voting and acceptance of answers on Stack Overflow, the accepted answer is also the most voted answer in most cases. In addition to that, there were two participants who said they looked at amount of comments for an answer too. Both of them explained that they preferred the answer with more comments since it means more information was provided.

In sum, individuals viewed various answer indicators because those can provide clues that showing the usefulness and informativeness of an answer. Former research suggested that votes and accepted answer are the determining factors to the quality of the answer (Gantayat et al., 2015). Results retrieved from this research also confirmed to that notion to some extent. Fifty-one participants described that they checked votes of an answer when evaluating knowledge on online Q&A sites. Thirteen participants narrated that they looked upon whether an answer was accepted. As a result, votes of an answer and accepted answer were the two most cited factors under category of answer indicators.

The indicators were peripheral factors that reflect the usefulness and informativeness of an answer. Site users can retrieve the information regarding usefulness of the answer contained in those factors without reading through the answer content itself. Comparing

with reading the answer content throughout, checking answer indicators is much easier task that individuals can employ. Because indicators are nearly pure numeric values, whereas the answer content usually contains one or more sentences, checking indicators is less time-consuming. Based on the facts mentioned above, utilizing such indicators as evaluation criteria could make information filtering process faster and easier.

Answer indicators was the most dominant factor that people employed in this research. To some levels, those indicators from online Q&A sites reflect the collective opinions from site users regarding the helpfulness of the answer. People look into answer indicators indicating that they believe in those indicators to some extent, which also means that they trust others' opinions to other extent. As illustrated earlier, when people adopt answer indicators for evaluating knowledge, a peripheral route is applied. The advantages of peripheral process embody less cognitive effort and less time consuming (Bhattacharjee & Sanford, 2006; Fadel et al., 2015). These advantages are attractive to people. Especially when they have to retrieve information for solving a problem with time constraints (Ratneshwar & Chaiken, 1991). In addition to that, according to dual process theory, peripheral process is often used for linked and quick thinking (Strannegård, von Haugwitz, Wessberg, & Balkenius, 2013). The purpose for individuals to use online Q&A is to find proper answer to their questions. Sometimes, there are plenty answer alternatives, which make information filtering process difficult for individuals. Answer indicators can help individuals to fulfill their objectives quickly and easily.

Lengthiness

From the interviews, participants said that they also pay attention to length of an answer. Some of them preferred long answer, whereas others prefer short answer. Participants who believed long answer is better explained that long answer usually contains richer information. However, participants who preferred short answer concerned about conciseness of an answer. They illustrated that concise answer is usually short too. Concise answer is helpful for extracting the main idea easily and rapidly, therefore, to solve the problem at hand quickly.

Based on the facts illustrated above, individuals check the length of the answer to retrieve peripheral cues regarding answer quality. Whether longer answer or shorter answer can result in better quality was not concluded from this research. Instead, it looks like individuals have different impressions regarding the length of the answer, and the impression varies depending on individuals' own knowledge and experiences. Length of answer is also validated in prior research as the factor that can affect the quality of an answer (Gantayat et al., 2015). Therefore, checking the length of an answer can assist individuals on evaluating the answer.

Empirism

When individuals apply empirism for evaluating knowledge on online Q&A sites, they need to read through the answer at first to retrieve the knowledge. Thus, it can be seen as central process. Comparing to read only the peripheral cues like votes of the answer, going through the answer content usually takes more time and cognitive effort. In the interview, many participants stated that the knowledge evaluation was actually based on their intuition. Hence, it was an intuitive processing that participants employed. As introduced in previous section, intuitive process has similar advantages than peripheral process. Intuitive process has advantage of being proceeded quickly without employing too much critical thinking. It can also process multiple pieces of information in parallel with no constraints for cognitive capacity or the number of information (Glöckner, 2010.).

Consequently, adopting empirism for knowledge evaluation is central route process but can take advantages from intuition process. This result looks a bit conflicting, but it is reasonable. Intuitive process is not necessarily based on peripheral cues although intuitive processing has similar advantages as peripheral processing does. Intuition processing is mainly used for integrating information and decision-making phase in knowledge evaluation process (Glöckner, 2010). Therefore, the individual can take both central route and intuitive processes for knowledge evaluation. This is very interesting findings from this research.

Feasibility

Feasibility is the second most dominant factor found in this research. It contains four sub-categories: similarity to own contexts, solution is tested to be working, complexity to apply the solution, and comments from others. The first three sub-categories judge the answer content itself. Therefore, those could be regarded as factors that are considered in central processing. Comparatively, comments from others is not part of the answer content itself, it is indirect factor that could reflect the quality of an answer to some levels. Thus, it can be seen as the factor that emerges from peripheral processing. Furthermore, opinions from others is regarded as heuristic cues as well in some other research (Chaiken & Maheswaran, 1994). Based on the facts illustrated above, when people consider feasibility of an answer, both central and peripheral routes might occur at the same time.

Readability

Writing skill is important to answer creators on Q&A websites. It helps them accurately articulate their opinions in the answer (McNamara, Crossley, & McCarthy, 2010). Good writing skill is an essence to convincingly communicate to others (Crowhurst, 1990). However, writing for web requires different skills than the traditional writing like writing for standard press release. Unlike traditional writing, writing for web is preferred to be more concise and enticing, so that readers can quickly find the specific information they need before moving to other content (Wise, 2005). Findings from this research also supported that notion as some participants have expressed that they preferred concise answer.

Writing styles are closely related to readability (Klare, 1963). Readable information can result in better comprehension, retention, and perseverance. (DuBay, 2007, p. 5). In the meantime, it can also improve reading speed and efficiency (Klare, 1963). Because of the benefits that readable information can bring, individuals tend to lean to more readable information during reading (Klare, 1963). Recent research also indicates that readability is especially important for web content because it heavily affects the understandability of the content (Ojha, Ismail, & Kuppusamy, 2018).

In this research, answer linguistic quality, concise answer, and answer formatting were found to be relevant to readability. In previous research, using correct grammar and spelling, using short sentences, and simple structured styles of writing were also verified to be useful for improving readability of the content (Dubay, 2007, pp. 1-2; Klare, 1963; Ojha et al., 2018). It can be seen that findings generated from this research also confirmed to earlier research regarding readability.

In order to check readability of an answer on online Q&A sites, individuals have to read the answer content itself. Evaluating the answer based on factors relevant to readability requires cognitive effort from individuals. For instance, individuals need to know the correct grammar before they can check its correctness. Only then individuals can compare

the grammar used in the sentence to the knowledge about the grammar that exists in their brains. Hence, readability related factors can trigger central route for processing information.

Source credibility

Source credibility is defined by Ohanian (1990) as a term that is commonly used to reflect the positive characteristics regarding the knowledge provider, and source credibility can affect knowledge seekers' acceptance of the knowledge. Source credibility is also a double dimensional concept that covers expertise and trustworthiness (Hovland, Janis, & Kelley, 1953).

One of the main factors "source credibility" that is generated from this research contains several child factors including: answerer credibility, external sources, and consistency. Since question answerer is the knowledge provider, therefore the question answerer is the source of the knowledge. For that reason, it makes sense that individuals check credibility of the answer provider. Besides, according to some participants' responses, they checked the credibility of the answerer to see the expertise of the question-answerer. External sources are used for checking whether the knowledge can be supported by someone or some other sources than the answer provider himself/herself. Participants illustrated in the interview that if there was a link to external source that supports the knowledge provided in the answer, then they regard this answer as trustful. Hence, external sources are used as evidence that supports trustworthiness of the knowledge. Consistency is used by individuals for checking trustworthiness of the knowledge as well. According to participants, they believed that the answer is more trustable if the answer is similar to several other answers that are given for the same question.

To conclude, answerer credibility, external sources, and consistency are closely related with trustworthiness and expertise dimensions. Since trustworthiness and expertise are two key factors that affect source credibility, thus the factors (answerer credibility, external sources, and consistency) that are retrieved from this research can be categorized to main factor "source credibility".

Source credibility also has been verified as main hint that individuals can employ when making a decision. It also has an effect on the individual's opinion change towards a piece of information (Manfredo & Bright, 1991.). Certain amount of research has indicated that source credibility is closely relevant to knowledge quality (Rieh & Danielson, 2007; Rieh & Belkin, 1998; Savolainen, 2011). Findings from this research confirmed to prior research by revealing that participants consider source credibility when evaluating an answer from online Q&A sites, wherein answer evaluating is the process where the individual checks the quality of the knowledge.

Petty and Cacioppo (1986, pp.160) suggest that source credibility is powerful simple cue that can affect individuals' judgement. That indicates source credibility can be regarded as peripheral cue to evaluate the knowledge. Indeed, checking source credibility does not take much cognitive effort from individuals, whereas it is still helpful for evaluating the knowledge. For instance, on Stack Overflow, the community users can earn reputation by posting good questions and useful answers. The reputation score of the answerer is shown besides profile image of the answerer, site users can easily check that numeric value, then quickly judge that whether the answerer is credible to some extent.

Recency

Participants valued the recency of an answer in evaluation process. According to them, recency can ensure the validity of the answer to some extent. Technologies from software engineering industry is changing rapidly and sometimes significantly (Boehm, 2006; Zhu, Shen, Cai, & Wang, 2015). That makes the software engineering related information is time-sensitive. Therefore, in cases where information timeliness is a necessity, it is helpful to check the recency of the answer.

When individuals view information, they need to continuously integrate the incoming information and transfer the information into comprehensive knowledge which is stored in their memory. If there is new information comes in, the new information will be combined, so that the individual's memory is updated (O'brien, Cook, & Guéraud, 2010.). Sometimes, not only the new information is combined, but also the outdated information might be abated (Johnson, & Seifert, 1998). For example, on Stack Overflow, the answer seeker goes through answers that were given for solving a framework problem. At first the answer seeker checks an answer that was marked as accepted answer two years ago. Then the answer seeker realized that since the framework version has been updated, the old solution from two years ago did not fit into the new version anymore. Therefore, the answer seeker checks another answer which was given recently and applicable for the correct version of the framework. As a result, the old answer was abandoned by answer seeker and the latest working answer was accepted.

Based on the interpretation explained above, the date of an answer can reflect the answer correctness and validity to some extent. The individual checks when the answer was given on online Q&A sites to retrieve the most updated information, therefore to acquire the valid knowledge.

To sum up, five out of eight main factors are verified as heuristic cues of an answer on online Q&A sites. A sub-category of feasibility, comments from others is also heuristic cue. This indicates that when site users evaluate answers on Stack Overflow, they lean a lot on heuristic cues than the content itself. Even though prior research has shown that when information seekers have low motivation and expertise, they rely more on peripheral cues (Fadel et al., 2015). However, there was no evidence found in this research which shows that participants fulfilled such characteristics. Instead, we suggested that the reason why site users rely on so many heuristic cues is that they are economy-minded in most cases (Fiske & Taylor, 1991, pp. 15-16). That means they try to retrieve the knowledge with least cognitive and time cost. Peripheral process can benefit in taking less time and cognitive effort (Bhattacharjee & Sanford, 2006; Fadel et al., 2015), hence, it results in quicker knowledge retrieval. That can explain why site users preferred more those heuristic cues than the answer content itself.

5.2 Interpretation based on information quality literatures

The research question of this study is to investigate how people evaluate knowledge acquired from online Q&A sites. Those categories appeared from interview can help answer the research question. Findings from this research can offer new outlook to knowledge evaluation, and also confirm a number of literatures regarding information quality.

In this research, it was found that many of factors confirm the dimensions of a quality model conducted from prior research (Zhu et al., 2009). The quality model contains thirteen dimensions, which includes: informativeness, politeness, completeness,

readability, relevance, conciseness, truthfulness, level of detail, originality, objectivity, novelty, usefulness and expertise. The quality model is made for assessing the user-generated content from some Q&A sites. Our research site is also a type of Q&A site, and we are investigating how people evaluate the answers on it. The two research contexts are similar. Thus, we believe it is reasonable to compare our findings with Zhu et al. (2009)'s quality model. In the meantime, we also hope to have a better understanding of our findings after comparison. The comparison of our findings to quality model is listed in Table 2.

To conclude, this research compiled well to the quality model. Most of the factors emerged from this research have relevant dimensions in the quality model (Zhu et al., 2009). That includes: living example, answer indicators, lengthiness, feasibility, readability, and source credibility. Findings from this research supported the quality model by indicating that informativeness, usefulness, conciseness, truthfulness, relevance, readability, and level of detail were also considered by individuals when they evaluate answers on online Q&A sites.

Two factors (empiricism and recency) emerged from this research are extra dimensions to that quality model. Nearly 36% of the participants from this research showed that they employ empiricism for knowledge evaluation on online Q&A sites. Empiricism has notable advantages of quicker evaluation process. Meanwhile, empiricism usually works unconsciously (Betsch & Glöckner, 2010). Based on that, empiricism is very likely to be employed by individuals for knowledge evaluation.

Recency is another factor that is supplementary to the quality model. Approximately 18% of participants consider recency as one metric for evaluating quality of answers. Its importance to individuals has been discussed in the previous section. In addition, recency is especially important to some online Q&A sites under domain like programming. Thus, recency is a valuable factor for knowledge evaluation.

Empiricism and recency are the factors that are complementary to the quality model. Both factors were verified as useful for information evaluation. The findings may provide hints for expanding dimensions of the quality model, therefore to make it more comprehensive.

In addition to the quality model, some of the main factors found from this research also confirm other literatures from the field of information quality.

The result of this research shows that amount of comments has an effect on individuals' knowledge evaluation process. In other research under the domain of online forum, number of replies is regarded as an indicator of post value, it also shows user's interest regarding the post (Wanas, El-Saban, Ashour, & Ammar, 2008). This research confirms that notion by reflecting that various answer indicators can affect individuals' evaluation regarding knowledge quality.

Surface features are defined as the way a post is displayed regardless of the actual content. Surface features are found to be relevant to post quality regardless of the post content. The reason why surface features can reflect post value is that it shows how careful the post writer writes the post. That could affect site users' perception regarding the post value (Wanas et al., 2008.). Lengthiness is one metric that can be used to assess surface features (Wanas et al., 2008, Weimer & Gurevych, 2007). In this research, length of answer is also found as one factor that could affect individuals' evaluation regarding knowledge quality.

Table 2. Comparison of our findings to information quality model

Main factor found from this research	Dimension from quality model	Explanation of why factor confirms to the certain dimension of quality model
Living example	Informativeness	Living example contains lots of information that is relevant to solve the problem.
	Usefulness	If the code example provided in the answer can actually help answer seeker solve the problem, then this code example is useful.
Answer indicators	Usefulness	Answer indicators (votes of answer, votes of comment, accepted answer) indicate how useful the answer is.
Lengthiness	Informativeness	Participants who think long answer is better often refer to “informativeness” dimension. They explained that longer answer usually contains richer information.
	Conciseness	Participants who prefer short answer concern about the “conciseness” dimension. They illustrated that concise answer is usually short too.
Feasibility	Truthfulness	There are four sub categories under feasibility in our findings, including: similarity to own contexts, solution is tested to be working, complexity to apply the solution, and comments from others. Solution is tested to be working and complexity to apply the solution are clearly checking how feasible the answer is. That confirms to the feasible dimension from truthfulness. Many participants illustrated that they checked the comments from others to look for evidence that convincing the answer is actually trustable. For this reason, comments from others are also for checking the truthfulness.
	Relevance	Similarity to own contexts is a sub-category from feasibility, it actually checks “relevance” dimension.
Readability	Readability	Answer linguistic quality and answer formatting refer to the readability dimension.
	Level of detail	Writing quality includes one sub-category named level of detail, which perfectly confirms to “level of detail” dimension from quality model.
	Conciseness	Concise answer sub-category concerns conciseness dimension.
Source credibility	Truthfulness	Source credibility is to check how trustable an answer is, which shows the truthfulness of an answer.

Another metric for assessing surface features is formatting quality. If a post is easy to read, readers will regard the post as with greater value (Wanas et al., 2008). Formatting quality includes aspects of using punctuations, emoticons, and capital letters (Wanas et al., 2008, Weimer & Gurevych, 2007). Answer formatting and answer linguistic quality are generated through this research as the child factors of readability that affect individuals' knowledge evaluation. The findings confirm the prior research appropriately.

In this research, source credibility is found to be useful for assessing the knowledge acquired from online Q&A sites, in which external sources is one metric. External sources could be a link to another site or page for instance. Taking advantage of the value from referenced source, web links are able to add value and credibility to a post (Wanas et al., 2008, Weimer & Gurevych, 2007). Our finding regarding source credibility reinforces that assertion.

6. Discussion

This research investigated individuals' knowledge evaluation process on online Q&A sites. Specifically, an exploratory study was conducted, in which sixty-seven participants of Stack Overflow were interviewed. Eight factors were revealed in this research, which are important for people to evaluate knowledge on online Q&A sites. The findings were interpreted based on dual process theory and information quality literatures. Both confirmation and novel standpoint to prior research were found. In addition, this research also has useful implications for practice.

6.1 Discussion of findings

Specifically, the eight main factors (that were found to be influential for site users in their knowledge evaluation process) are code example, answer indicators, lengthiness, empirism, feasibility, readability, source credibility, and recency. Among all factors, answer indicators and feasibility are the most dominant. Living example and empirism follow after. Whereas lengthiness is the least popular factor, even though it has approximate points as readability, source credibility and recency have.

Findings from this research reflect that when individuals evaluate knowledge retrieved from online Q&A sites, they employ a dual process which contains central and peripheral routes. Some factors that can affect knowledge evaluation process can trigger central routes, which include code example, empirism and feasibility. Comparatively, other factors are heuristic cues of an answer in essence, and thus these factors can trigger the peripheral route. The peripheral factors include answer indicators, lengthiness, readability, source credibility, recency, and comments from others (sub-category of feasibility).

Central route processing and peripheral route processing occur interactively in individuals' knowledge evaluation process. The results from this research reflect that individuals incline towards evaluating the knowledge acquired from online Q&A sites by peripheral cues of the answer.

6.2 Contributions to the literature

In this research, dual process theory is applied to explain individuals' knowledge evaluation process on online Q&A sites. A dual process is found to be employed by individuals during the knowledge evaluation process on online Q&A sites, in which they assess not only the answer content itself, but also heavily rely on the heuristic cues of the answer. This is novel and paves the way for the research on knowledge evaluation over the Internet. Some factors (answer indicators, lengthiness, readability, source credibility, and recency) generated through this research refer to the peripheral clues of an answer, whereas some others (living example, empirism, and feasibility) are about the answer content itself. The facts retrieved through this research indicate that individuals check not only the answer content itself, but also the heuristic cues of the answer. Through applying knowledge from dual process to this research, the research phenomenon is comprehensively understood. The findings not only answered the research question, but also provided a new perspective regarding how similar research can be investigated.

Apart from that, the findings from this research also add new knowledge to literatures regarding information quality by indicating that empirism and recency are two dimensions that could be added to information quality model. Moreover, the findings also provide

evidence that supports the former literatures in context of information quality. Most of the factors generated through this research confirm the dimensions from quality model for assessing information quality. Besides, answer indicators, lengthiness, readability and source credibility were verified to be the indicators of answer quality on online forums. Therefore, individuals can take those factors into consideration to evaluate the knowledge on online Q&A sites.

Individuals' knowledge evaluation process is studied through revealing the factors that individuals consider to assess the answer quality. The findings from this research is supplementary to the research in context of knowledge evaluation, information quality, and online Q&A sites in particular, and can enrich the dual process theory in general.

6.3 Implications for practice

This research has extremely practical implications to practitioners on online Q&A sites. There are two main implications found, one for site managers, the other for site users.

Firstly, the findings have practical implications for site managers of online Q&A sites such as Stack Overflow. In this research, individuals' evaluation process in knowledge retrieving was studied. By understanding the process, site managers could get to know better the site users' criteria and attitude towards the answer quality. With that knowledge, site managers can take actions to provide a more friendly user experience. For instance, by emphasizing the more dominant factors that most individuals consider, site users can find those factors easily and immediately.

Secondly, findings from this research are also helpful to answer writers on guiding them writing more preferred answer for fellow users. For example, from the findings, many individuals favor answers with code example since those are more visualized. Answer writers could take that into account by adding a code snippet into their answer, so that their answers would be more possibly to attract site users' eyes.

To summarize, understanding individuals' evaluation process regarding knowledge retrieved from online Q&A sites is essential for the contemporary development of knowledge sharing community. It can not only guide knowledge seekers to acquire knowledge more efficiently, but also shed lights on managing the community of online Q&A sites to make them healthier and more user friendly. However, more research regarding individuals' knowledge evaluation process on the Internet environment should be conducted with certainty. More online Q&A sites with different domain than Stack Overflow could be studied to extend the relevant literatures.

6.4 Limitations and future research

In spite of the contributions this research could bring, some limitations also exist. One of the limitations is that the data collected for this research might lacked diversity. Most of the participants located in Finland and China. Thus, the results might not be generalized enough. It would be helpful to expand data diversity by having participants from other countries too.

In this research, evaluation process employed by site users on Stack Overflow was investigated. After analyzing all the data collected from interviews, a summary of factors that individuals consider was generated. However, each individual's evaluation process is not analyzed, and is not compared with others either. By doing that, however, researchers can have a more comprehensive understanding regarding the knowledge

evaluation process. Additionally, it would be interesting to see how individuals' knowledge evaluation process could differ from each other, and what would be the reasons to explain that.

Stack Overflow is a programming online Q&A site. That might have effect on our results to be applicable in knowledge evaluation research in general. There might be some specific factors that people would consider only in the context of Stack Overflow, or in general the programming online Q&A sites. For example, the living example factor mainly refer to code example in context of Stack Overflow, therefore, whether the living example from other online Q&A sites would affect individuals' knowledge evaluation process the same way as it does for Stack Overflow users remains uncertain. For that reason, we encourage researchers who are in interested can conduct studies to approach how the domain of the online Q&A sites would affect people's knowledge evaluation process.

This research has proposed that when individuals evaluate knowledge on online Q&A sites, both central and peripheral routes were employed. But there was no evidence showing the shift between central and peripheral routes from individuals. Therefore, it could not confirm the work from Fadel et al. (2015), which asserted that individuals can freely shift between central and peripheral routes depends on the ongoing evaluation task. Future research is suggested to have an experiment which tracking individuals' shift between central and peripheral process.

This research generated a summary of popularity of each factor that is influential to knowledge evaluation process on online Q&A sites. However, the popularity of each factor may not reflect the accuracy or quality of knowledge evaluation. It is reasonable to conduct a future research, which should somehow investigate the effect of each factor on knowledge evaluation accuracy. Understanding the effect on knowledge evaluation accuracy for each factor can guide knowledge seekers on more accurate knowledge evaluation and more efficient knowledge retrieval.

7. Conclusion

Online Q&A sites evolve to be popular online knowledge repository for knowledge seekers. With the prompt development of online Q&A sites, plentiful knowledge was created and stored on online Q&A sites. That enlarges the online knowledge repository, whereas brings obstacles for site users at the same time. Investigating individuals' knowledge evaluation process on online Q&A sites is needed and beneficial.

This research was conducted to address the question of how people evaluate knowledge acquired from online Q&A sites. Stack Overflow was used as the research site because of its popularity. In total, sixty-seven Stack Overflow participants were interviewed in this qualitative research. Eight factors were generated through data analysis, including: living example, answer indicators, lengthiness, empirism, feasibility, readability, source credibility and recency. Answer indicators and feasibility are the most dominant factors.

A dual process is found to be employed by individuals during the knowledge evaluation process on online Q&A sites, in which they assess not only the answer content (living example, empirism, and feasibility) itself, but also heavily rely on the heuristic cues (answer indicators, lengthiness, readability, source credibility, and recency) of the answer. The findings pave the way for the research on knowledge evaluation over the Internet. The findings also enrich the information quality literatures. As a field study, this research has important implications for practice. Given the popularity of online Q&A sites, the importance of knowledge sharing, and the difficulty in knowledge evaluation, such research is certainly valuable.

References

- About - Stack Overflow. (2018). Retrieved March 15, 2018, from Stack Overflow: <https://stackoverflow.com/company>
- Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008, February). Finding high-quality content in social media. In *Proceedings of the 2008 international conference on web search and data mining* (pp. 183-194). ACM.
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2012, August). Discovering value from community activity on focused question answering sites: a case study of stack overflow. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 850-858). ACM.
- Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109.
- Betsch, T. (2008). The nature of intuition and its neglect in research on judgment and decision making. *Intuition in judgment and decision making*, 3-22.
- Betsch, T., & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry*, *21*(4), 279-294.
- Bhattacharjee, A., & Sanford, C. (2006). Influence processes for information technology acceptance: An elaboration likelihood model. *MIS quarterly*, 805-825.
- Boehm, B. (2006, May). A view of 20th and 21st century software engineering. In *Proceedings of the 28th international conference on Software engineering* (pp. 12-29). ACM.
- Boyd, C. (2017, November 22). WeChat: The evolution and future of China's most popular app. Medium. Retrieved March 18, 2018, from <https://medium.com/swlh/wechat-the-evolution-and-future-of-chinas-most-popular-app-11effa5639ed>
- Brenner, M., Brown, J., & Canter, D. (1985). *The research interview, uses and approaches*. Academic Press.
- Brinkley, M., & Burke, M. (1995). Information retrieval from the Internet: an evaluation of the tools. *Internet Research*, *5*(3), 3-10.
- Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of personality and social psychology*, *66*, 460-460.
- Chen, S., Duckworth, K., & Chaiken, S. (1999). Motivated Heuristic and Systematic Processing. *Psychological Inquiry*, *10*(1), 44-49.

- Crowhurst, M. (1990). Reading/writing relationships: An intervention study. *Canadian Journal of Education/Revue canadienne de l'éducation*, 15(2), 155-172.
- Dictionary, f. (2018). *feasibility* Meaning in the Cambridge English Dictionary. Dictionary.cambridge.org. Retrieved March 2, 2018, from: <https://dictionary.cambridge.org/dictionary/english/feasibility>
- DuBay, W. H. (2007). *Smart language: Readers, readability, and the grading of text*. Costa Mesa, CA: Impact Information.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, 62(1), 107-115.
- Evans, J. S. B. (2003). In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences*, 7(10), 454-459.
- Fadel, K. J., Meservy, T. O., & Jensen, M. L. (2015). Exploring knowledge filtering processes in electronic networks of practice. *Journal of Management Information Systems*, 31(4), 158-181.
- Fiske, S. T., & Taylor, S. E. (1991). *McGraw-Hill series in social psychology. Social cognition (2nd ed.)*. New York, NY, England: Mcgraw-Hill Book Company.
- Gantayat, N., Dhoolia, P., Padhye, R., Mani, S., & Sinha, V. S. (2015, May). The synergy between voting and acceptance of answers on stackoverflow, or the lack thereof. In *Proceedings of the 12th Working Conference on Mining Software Repositories* (pp. 406-409). IEEE Press.
- Ge, M., & Helfert, M. (2006). A review of information quality research. In *11th International Conference on Information Quality* (pp. 159-172).
- Goldstein, D. G., & Gigerenzer, G. (2011). Models of Ecological Rationality: The Recognition Heuristic. *Heuristics*, 60-80.
- Gyongyi, Z., Koutrika, G., Pedersen, J., & Garcia-Molina, H. (2007). *Questioning yahoo! answers*. Stanford InfoLab.
- Heiskanen, A., & Newman, M. (1997, December). Bridging the gap between information systems research and practice: the reflective practitioner as a researcher. In *Proceedings of the eighteenth international conference on Information systems* (pp. 121-132). Association for Information Systems.
- Hoepfl, M. C. (1997). Choosing Qualitative Research: A Primer for Technology Education Researchers. *Journal of Technology Education*, 9(1), 47-63.
- Honsel, V., Herbold, S., & Grabowski, J. (2015, May). Intuition vs. truth: Evaluation of common myths about stackoverflow posts. In *Mining Software Repositories (MSR), 2015 IEEE/ACM 12th Working Conference on* (pp. 438-441). IEEE.

- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; psychological studies of opinion change*. New Haven, CT, US: Yale University Press.
- Imqq.com. (2018). *QQ International - Chat, Video Calls, Groups | Get a Better Messenger*. Retrieved March 2, 2018, from: <http://www.imqq.com/English1033.html>
- James, N., & Busher, H. (2012). Internet interviewing. *The SAGE handbook of interview research: The complexity of the craft*, 177-188.
- Johnson, H. M., & Seifert, C. M. (1998). Updating accounts following a correction of misinformation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6), 1483.
- Klare, G. Roger. (1963). *The measurement of readability*. Ames: Iowa State University Press.
- Liu, Y., Bian, J., & Agichtein, E. (2008, July). Predicting information seeker satisfaction in community question answering. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 483-490). ACM.
- Mack, N., & Woodson, C. (2005). *Qualitative research methods: A data collectors field guide*. North Carolina: FLI.
- Manfredo, M. J., & Bright, A. D. (1991). A model for assessing the effects of communication on recreationists. *Journal of Leisure research*, 23(1), 1-20.
- Mattermost. (2018). *Mattermost Private Cloud Messaging*. Retrieved March 2, 2018, from: <https://about.mattermost.com/>
- McNamara, C. (2009). General Guidelines for Conducting Research Interviews. Retrieved January 30, 2018, from <https://managementhelp.org/businessresearch/interviews.htm#anchor1404957>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written communication*, 27(1), 57-86.
- Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and organization*, 17(1), 2-26.
- Nasehi, S. M., Sillito, J., Maurer, F., & Burns, C. (2012, September). What makes a good code example? A study of programming Q&A in StackOverflow. In *Software Maintenance (ICSM), 2012 28th IEEE International Conference on* (pp. 25-34). IEEE.
- O'brien, E. J., Cook, A. E., & Guéraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 979.
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of advertising*, 19(3), 39-52.

- Ojha, P. K., Ismail, A., & Kuppusamy, K. S. (2018). Perusal of Readability with Focus on Web Content Understandability. *Journal of King Saud University-Computer and Information Sciences*.
- Oliver, E. (2007, Jul). Intuition in decision-making. *Nonprofit World*, 25, 11-13. Retrieved from <https://search.proquest.com/docview/221336913?accountid=13031>
- Opendakker, R. (2006, September). Advantages and disadvantages of four interview techniques in qualitative research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (Vol. 7, No. 4).
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications, inc.
- Petty, R. E., & Cacioppo, J. T. (1986). *Advances In Experimental Social Psychology* (Vol. 19). Academic: Press.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218.
- Powell, S. R., & Hebert, M. A. (2016). Influence of writing ability and computation skill on mathematics writing. *The Elementary School Journal*, 117(2), 310-335.
- Ratneshwar, S., & Chaiken, S. (1991). Comprehensions Role in Persuasion: The Case of Its Moderating Effect on the Persuasive Impact of Source Cues. *Journal of Consumer Research*, 18(1), 52-62.
- Rew, L. (1988). Intuition in Decision-making. *Journal of Nursing Scholarship*, 20(3), 150-154.
- Rheingold, H. (1994). *The virtual community: Finding connection in a computerised world*. London, England: Secker & Warburg.
- Rieh, S. Y., & Belkin, N. J. (1998, October). Understanding judgment of information quality and cognitive authority in the WWW. In *Proceedings of the 61st annual meeting of the american society for information science* (Vol. 35, pp. 279-289).
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual review of information science and technology*, 41(1), 307-364.
- Savolainen, R. (2011). Judging the quality and credibility of information in Internet discussion forums. *Journal of the Association for Information Science and Technology*, 62(7), 1243-1256.
- Shah, C., & Pomerantz, J. (2010, July). Evaluating and predicting answer quality in community QA. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 411-418). ACM.

- Stack Overflow - Where Developers Learn, Share, & Build Careers. (2018). Retrieved January 26, 2018, from Stack Overflow <https://stackoverflow.com/>
- Strannegård, C., von Haugwitz, R., Wessberg, J., & Balkenius, C. (2013, July). A cognitive architecture based on dual process theory. In *International Conference on Artificial General Intelligence* (pp. 140-149). Springer, Berlin, Heidelberg.
- Turner, III, D. W. (2010). Qualitative interview design: A practical guide for novice investigators. *The qualitative report*, 15(3), 754.
- Wanas, N., El-Saban, M., Ashour, H., & Ammar, W. (2008, October). Automatic scoring of online discussion posts. In *Proceedings of the 2nd ACM workshop on Information credibility on the web* (pp. 19-26). ACM.
- WeChat. (2014). *Tencent Inc.* [App]. Retrieved March 2, 2018, from <https://itunes.apple.com/fi/app/wechat/id836500024?mt=12>
- Weimer, M., & Gurevych, I. (2007, September). Predicting the perceived quality of web forum posts. In *Proceedings of the conference on recent advances in natural language processing* (pp. 643-648).
- Wise, K. (2005). The importance of writing skills. *Public Relations Quarterly*, 50(2), 37.
- Zhu, J., Shen, B., Cai, X., & Wang, H. (2015). Building a Large-scale Software Programming Taxonomy from Stackoverflow. In *SEKE* (pp. 391-396).
- Zhu, Z., Bernhard, D., & Gurevych, I. (2009). *A multi-dimensional model for assessing the quality of answers in social Q&A sites* (Doctoral dissertation).
- Zwass, V. (2017). CM: Information System. Retrieved January 16, 2018, from <https://global.britannica.com/topic/information-system>