It's Not What You Say, It's What You Do: The Motivation of The Crowd to Participate in a Crowdsourcing Project to Support Blind and Partially Sighted Students

Fatma Layas

Doctor of Philosophy

University of York
Computer Science
June 2017

Abstract

There is a growing interest in crowdsourcing projects for socially responsible issues. One area of socially responsible crowdsourcing is to support people with disabilities. However, there is little exploration of what motivates people to participate in such projects. This programme of research investigated the motivators for students to participate in a socially responsible crowdsourcing project to support blind and partially sighted students by describing images found in digital learning resources. For this purpose a crowdsourcing project, DescribeIT, was developed.

The first study explored what students thought would motivate them to participate in the project to compare with students' actual behaviour in the following studies. Altruism and monetary rewards were the leading self-reported motivational factors, other factors such as being interested in accessibility were reported.

Studies 2 to 6 investigated the effects of different intrinsic and extrinsic motivational factors on students' participation in the DescribeIT project with students from the UK and Arab countries. Despite the promising results of the self-reports of motivations, UK students' participation rates in Studies 2 to 4 was extremely low. However, paying UK students small amounts of money (Study 6) did motivate them to participate. Arab students (Study 5) were intrinsically motivated to participate in the DescribeIT project and showed a higher participation rate than UK students.

Studies 7 and 9 investigated the quality of the image descriptions produced by crowd members of established crowdsourcing platforms in comparison to those produced by students. The results showed a comparable quality across descriptions produced by students and crowd members.

Studies 8 and 9 investigated the effect of simplifying the image description task by changing it to an image tagging task and showed that making the task easier increased participation rate. Lastly, Study 10 investigated the effect of a face-to-face training session on image description quality. It also investigated the effect of quality control instructions on quality. The face-to-face training increased description quality, but different quality control instructions did not.

The practical implications of this research for crowdsourcers in socially responsible crowdsourcing contexts, are that they need to consider the cultural backgrounds of their potential crowd, make the task easy to do, offer small payments if possible and train crowd members in order to produce good quality work. The theoretical implications are a greater understanding of the motivations of crowd members in socially responsible projects and the importance of measuring both self-reports of motivation and actual behaviour.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to express my gratitude to Allah the Almighty for giving me the power and blessing to complete my studies.

This thesis would not have been possible without the help and support of many people, and I would like to take this opportunity to extend my sincere gratitude and thanks to all of them.

I owe my deepest gratitude to a truly brilliant mind my supervisor Professor Helen Petrie. Thank you, Helen, for the untiring support, help, patience, and encouragement throughout my studies, which made this work possible. Your advice and guidance have been priceless. I have been extremely lucky to have you as my supervisor. Thank you from the bottom of my heart for being there for me whenever needed in my research and beyond.

I would also like to thank Dr. Christopher Power for all the fruitful research discussions we had. Thank you, Chris!

I would like to express my gratitude for my internal examiner Dr. Paul Cairns for his insightful comments and challenging questions in our TAP meetings. Thank you, Paul, for providing me with a second view on my work.

I am also grateful to Anna Bramwell-Dicks, David Zendle, Dr. William Smith, and Dr Alistair Edwards for helping in recruiting participants for my studies. In addition, I would like to thank all the numerous people who took part in my studies.

In addition, I would like to thank the Ministry of Higher Education and Scientific Research in Libya for funding my PhD studies.

I am immensely grateful to Andreas Savva for his valuable assistance in developing the project used in my programme of research. A special thanks to my dear and special friend Saja Althubaiti who was always there for me, during the ups and downs through this entire process. My thanks and appreciations also go to all my colleagues whom I am happy to call my friends, David Swallow, Zaidatol Haslinda Abdullah, Leonardo Sandoval Guzman, Burak Merdenyan, Sarah Alotaibi and Maneerut Chatrangs. Thank you guys so very much for everything!

I would like to thank the entire Department of Computer Science for their support and help. Special thanks to, Jayne Fitzgerald, Helen Ducker, Virginie Harrison and Alex

Cooper whom were very helpful whenever needed. Also, I would like to thank Jenny and Yannis Darzentas for their kind words and encouragement, I am so happy that I had the chance to know both of you.

Thank you to all the lecturers who provided me with the digital teaching materials used in my studies from range of department in the University of York.

I also would like to make a special reference to whom without their generous support I would not be able to focus on my studies because of the situation in my country, the open-door team at the University of York, thank you all.

Words are inadequate to express my thanks to my husband Ahmad Benkhial for his unconditional love, tremendous support, and providing me with strength during the most difficult times in this journey.

I would like to express my heartfelt thanks to my beloved family for their blessings, endless support, encouragement and love through this journey and my life in general, especially my parents Abdul Monsef Layas and Halima Benkhial, who encouraged me to pursue this PhD in the first place. Thank you, dad for believing in me. Thank you mum for being my rock. I extend my thanks to my brothers Mohammed (and his wife), Marwan, Osama and Yosef Layas who always supported me in so many ways. I cannot image my life without you guys.

Words cannot express how grateful I am to my in-law family Uncle Abdul Hamed Benkhial, aunt Muna Sahli, Osama Benkhial, Arwa Benkhial for their enthusiastic support and prayers.

To my nieces and nephews Marwa, Abdul Allah, Safa, Maryam, Yazan, and Lareen I love you all so much. You bring so much joy into my life.

I also express my sincere thanks to my extend family for their good wishes and prayers.

I would also like to thank very special friends whom I consider family aunt Aisha, Asma, Misoon, Ebtisam, Najwa, and Iman who I had the extraordinary privilege to get to know in York, and who opened their hearts before their doors for me and my husband. Ladies you will always have a special place in my heart.

At the end, I would like to offer my regards and thanks to all the people who were willing to help me with their abilities in this journey.

Declaration

I declare that the work presented in this thesis, if not otherwise stated, is my own. Some of the material presented within this thesis has previously been published in the following conferences:

Layas, F., Petrie, H., & Power, C. (2015, September). A Cross-Cultural Study of Motivations to Participate in a Crowdsourcing Project to Support People with Disabilities. In Third AAAI Conference on Human Computation and Crowdsourcing.

Layas, F., Petrie, H., & Power, C. (2015, October). A Study of Students' Motivations to Participate in a Crowdsourcing Project to Support their Disabled Peers. In Eighth York Doctoral Symposium on Computer Science & Electronics (p. 47).

Layas, F., & Petrie, H. (2016). An Assessment of Motivational Factors on Participation and Performance in Socially Responsible Crowdsourcing Projects. International conference on Business, Economics, Social Science & Humanities, volume 247, issue 29, (p. 13)

Layas, F., & Petrie, H. (2016). Exploring Intrinsic and Extrinsic Motivations to Participate in a Crowdsourcing Project to Support Blind and Partially Sighted Students. Studies in health technology and informatics, 229, 545-555.

Petrie, H., & Layas, F., (2017) Crowdsourcing the descriptions of images in interactive digital media. Accessibility in Film, Television, and Interactive Media Conference. *Accepted and to appear in the conference proceeding.*

To my parents Abdul-Monsef and Halima

&

My husband Ahmad

Chapter 1 Introduction

Crowdsourcing is *the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call* (Howe, 2006). It has been used in many areas, both in the private sector and for socially responsible purposes. One socially responsible area is to support people with disabilities. However, there is little exploration of what motivates people to participate in this type of crowdsourcing projects.

Many researchers have proposed projects to harness the knowledge of the crowd (the term used to refer to people who participate in crowdsourcing) by using crowdsourcing to be a useful means of undertaking socially responsible activities, such as providing access to online information for people with disabilities. There have been various projects applied to different aspects of accessibility such as the creation of accessible maps (Cardonha et al., 2013), documenting where sidewalks (Hara, Le, and Froehlich, 2012) and bus stops (Hara et al., 2013) are accessible, describing photos taken by blind users to them (Bigham, et al. 2010), tagging images (Von Ahn and Dabbish 2004), and fixing web accessibility problems (Takagi, Kawanaka, Kobayashi, Itoh and Asakawa, 2008). These initiatives have shown that crowdsourcing has the potential to be a useful means of undertaking socially responsible activities. However, many technical and practical challenges need to be addressed to achieve a successful crowdsourcing project. One of those challenges is the need to use appropriate mechanisms to ensure that people are motivated to participate in such crowdsourcing projects.

The success and sustainability of many crowdsourcing projects such as Wikipedia, is largely dependent on the willingness of people to contribute and share their knowledge, time, effort and skills voluntarily in contributing to these projects. Understanding the motivations and willingness of the crowd members to participate in crowdsourcing projects is receiving a great deal of attention from researchers. However, an open question of investigation is the key motivators that will ensure people’s participation in socially responsible crowdsourcing projects. A number of studies have investigated crowd members’ motivations in a range of contexts, including micro-payment tasks (Kaufmann, Schulze and Veit, 2011), artistic designs (Brabham 2008, 2010, 2012), and contests (Zheng, et al. 2011). However, little of the research investigated the motivations of the crowd in the context of socially responsible crowdsourcing.

With the rapid development of the use of technology in education, higher educational institutions now often use virtual learning environments (VLEs) to facilitate teaching and learning. Lecturers typically upload their teaching materials, which can include PowerPoint slide packs used for lectures and seminars. Students can study these before the lecture or seminar, and review them afterwards. While there are many types of digital learning resources in higher education settings that require adaptation for people with disabilities, this programme of research will use the provision of descriptions of images in digital learning resources for people with visual disabilities as an example of crowdsourcing task that people could undertake in an educational setting. The task of providing image descriptions was chosen as it is a long-term and on-going problem for blind and partially sighted students in many educational settings (Splendiani and Ribera, 2014; Splendiani, Ribera, Centelles Velilla, 2014). When one considers digital learning resources, such as PowerPoint presentations, these include many images, such as: photographs, graphs, diagrams and drawings. Which are very popular in digital learning resources, it is important to address the accessibility problems of these materials (Raskind, et al, 2005). If these are not explicitly described, visually disabled student may be severely disadvantaged in their learning. However, lecturers rarely provide such descriptions, and lack the time to do so. The fact that the number of disabled students is increasing (Higher Education Statistical Agency, 2013) means that the number of students who are affected by inaccessible digital learning resources is increasing. In this context, providing image descriptions for images in such resources will help improve the educational experience for blind and partially sighted students.

1.1 Research questions

The key research question of this programme of research is:

What are the motivators for students to participate in socially responsible crowdsourcing projects?

The research question was investigated in one particular domain, that of non-disabled students participating in a crowdsourcing project to create descriptions of images in digital learning resources for visually disabled students. A research programme based around such a crowdsourcing project for creating image descriptions has the potential to both answer the above research question, while also providing an immediate benefit to visually disabled students in terms of increasing the accessibility of digital learning resources. The key research question was broken down to a number of more specific research questions, each study or a group of studies in this programme of research addresses these as follows:

What do students think would motivate them to participate in a socially responsible crowdsourcing project? [Study 1 – Chapter 4]

Does students' self-reported motivation match their actual behaviour in the socially responsible crowdsourcing project? [Studies 2, 3 and 4, Chapter 5; Study 5 – Chapter 6]

Does monetary reward increase students' participation in the socially responsible crowdsourcing project? [Studies 6 and 7 – Chapter 7]

Does the difficulty of the task affect students' (and others) motivation? [Studies 8 and 9 – Chapter 8]

Does training improve students' performance in the image description task? [Study 10 – Chapter 9]

How does motivation influence students' (and others) participation in the socially responsible crowdsourcing project? [Studies 5 to 10]

How does students' (and others) sense of altruism influence their participation in the socially responsible crowdsourcing project? [Studies 5 to 10]

How does students' (and others) attitude toward people with disability influence their participation in the socially responsible crowdsourcing project? [Studies 5 to 10]

What motivate students from different culture background to participate in a socially responsible crowdsourcing project? [Study 5 – Chapter 6]

1.2 Research Approach

In order to investigate what are the key motivators that would motivate students to participate in a crowdsourcing project to support students with visual disabilities, a crowdsourcing project called *DescribeIT* was developed. The project allows sighted students (or others) to describe images in digital learning resources for blind and partially sighted students. The initial approach taken by this programme of research was to incorporate some of the motivational factors reported by potential users of the DescribeIT project into the project. This allowed the researcher to identify the motivational factors that do actually motivate students to participate in the DescribeIT project.

To this end, studies were conducted which investigated which motivational factors would influence students' participation rate, number of images they described and the quality of those descriptions.

The programme of research started with an exploratory study to investigate students' perceptions of what they thought would motivate them to participate in the DescribeIT project to support students with visual disabilities. This study established a baseline of self-reports of motivations to compare with actual behaviour in the project obtained in studies conducted subsequently in the thesis. This exploratory study was followed by Studies 2, 3, 4 and 5, which investigated how different motivational factors affect the participation of students in the socially responsible crowdsourcing project, DescribeIT. In this series of studies intrinsic and non-financial extrinsic motivational factors were manipulated. Each study consisted of two phases (see Figure 1.1), as follows: the first phase was a self-report of motivation to investigate what students thought would motivate them to participate in the DescribeIT project. To select which motivational factors to use in Phase 2, the motivational factors identified in Phase 1 were taken in consideration, as well as ones identified in the previous studies. For example, in Study 2 Phase 2, the results of Study 1 and Study 2 Phase 1 were taken in consideration. The second phase of each study investigated students' actual behaviour while using the DescribeIT project. A comparison between the results of Phase 1 and 2 was also made to investigate if there were differences between students' self-reports of their motivation and their actual behaviour.

The set of all possible motivational factors is too extensive to include in any one study. Therefore, each of the studies focused on one intrinsic and one extrinsic motivational factor. The motivational factors chosen were those which appeared to have the strongest motivational effect, based on students' self-reports. The set of motivational factors used in Studies 2 to 5 are summarized in in Table 1.1.

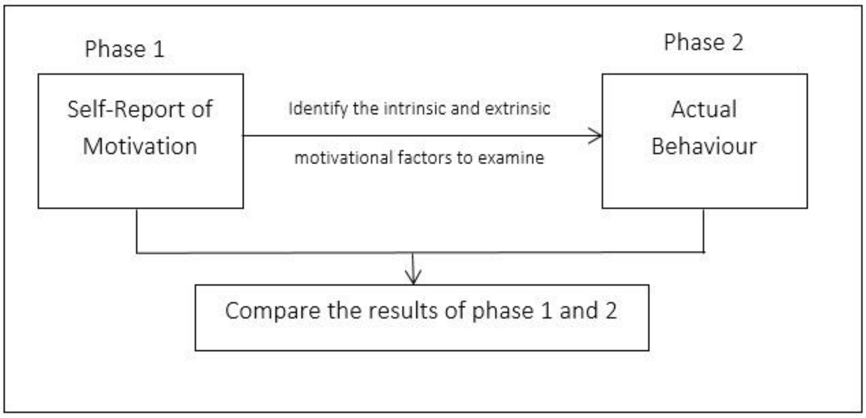


Figure 1.1 Design of Studies 2 to 5

Table 1.1 Motivational factors investigated in this programme of research

Motivational factor	Study
Altruism, wanting to help others	Study 2 (UK students, with some international students) Study 5 (Libya/Saudi Arabian students)
Improving one's skills	Study 2 (UK students, with some international students) Study 5(Libya/Saudi Arabian students)
Knowing that blind and partially sighted students will benefit from the project	Study 3 (UK students, with some international students)
Increasing students' understanding of the teaching materials	Study 3 (UK students, with some international students)
The effect of positive environment on participation	Study 4 (UK students, with some international students)
Financial rewards	Study 6 (UK students, with some international students) Study 7 (active crowd members)
Improving one's skills (Difficulty of the task was also investigated in this study)	Study 8 (UK students, with some international students) Study 9 (active crowd members)
Instructions about quality control of the descriptions (the effect of training in the image description task was also investigated in this study)	Study 10 (UK students, with some international students)

As mentioned earlier, the DescribeIT project aimed to support blind and partially sighted students in higher education by the description of images in digital learning resources. For this reason, university students were recruited to participate in the studies, often students doing the course from which the digital learning resources were taken, because they would understand the context and educational intent of the images. Students were recruited from a European country (UK) and from Arabic countries (Libya and Saudi Arabia). However, despite the promising results found in Study 1 and Phase 1 of Studies 2 to 5, students' participation rate in Phase 2 of Studies 2 to 4 ranged from a total lack of participation to a very low participation rate. Hence, it was decided to:

(1) Invite students who participated in Studies 2 to 4 to participate in Study 6, to investigated the effect of the important extrinsic motivational factor of monetary reward, on their participation in the DescribeIT project. The level of participation was then compared to the effects of the non-monetary motivational factors used in Studies 2, 3 and 4.
(2) Expand the sample population to include active crowd members (people who participate in crowdsourcing) from established crowdsourcing platforms. Therefore, in Studies 7 and 9 people were recruited from two different crowdsourcing platforms: Amazon Mechanical Turk and Crowdcrafting.org. The Amazon Mechanical Turk platform is a micropayment based platform (i.e. people are paid small amounts of money to do small tasks). Whereas, Crowdcrafting.org is a volunteer-based platform oriented to scientific research (i.e. people voluntarily do tasks).
(3) In Study 8, the task of describing images was made simpler by asking students to tag individual objects within the images instead of describing the whole image. The aim was to investigate whether the task that had been used in the previous studies in this programme of research, that of describing images, was perceived as too hard and off-putting for students or others.

The quality of the descriptions produced in each study was assessed as another measure of students' participation in the task. In addition, Study 10 investigated whether providing students with face-to-face training increased the quality of the descriptions produced. The question of students' perception of the difficulty of the image description task was also investigated in this study. Finally, this study investigated the effect of different messages in the instructions about how the quality of the descriptions produced would be assessed.

This programme of research investigated the varied types of motivational factors based on the motivational categories of Self Determination Theory (SDT) (Deci and Ryan 1985, 2000). SDT includes: (1) Intrinsic Motivation; (2) Extrinsic Motivation, including the sub-categories of Integrated Regulation, Regulation through Identification, Introjection Regulation, and External Regulation; and (3) Amotivation (See Chapter 2, section 2.4.2). The decision to use this theory was taken as it is one of the most popular and accepted theories of motivation. In addition, it has been used by many researchers (e.g. Kaufmann et al., 2011; Zheng et al., 2011; Brabham, 2012) to understand motivation in the context of crowdsourcing.

To measure the different types of motivation identified by SDT, it was decided to use the Situational Motivation Scale (SIMS) (Guay et al. 2000) (See Chapter 2, section 2.8 for more details about the scale). This decision was based on a number of reasons including:

1. The SIMS measures intrinsic motivation, extrinsic motivation (including two of the sub-categories: Identified Regulation and External Regulation) and Amotivation.
2. The SIMS measures the motivation of participants in a situational level, meaning measures participants' motivation while they are engaged with the image description task.
3. The SIMS is not limited to being used in lab settings, it can be used in real life settings.
4. The SIMS has been used to measure students' motivation previously, so it is suitable for the target group of this programme of research
5. The SIMS has acceptable levels of validity and reliability.
6. The SIMS is free to use.

To investigate whether people's general levels of altruistic behaviour and attitudes toward people with disabilities influence participation, it was decided to measure the former with the Self-Report Altruism Scale (SRA) (Rushton et al. 1981) and the latter with the Interaction with Disabled Persons Scale (IDP) (Gething 1994) (See Chapter 2, section 2.8 for more details about each scales). The decision to use these scales was based on a number of reasons including:

1. Even though both scales are quite old, both are still being used by many researchers (e.g. SRA Scale: Anderson and Clark, 2015; Arnocky et al., 2017; IDP Scale: Brown et al, 2009; Smith et al, 2014).
2. Both scales have been used with students in previous research (Rushton et al., 1981; Brown et al., 2009) which means that they are suitable for the target user group and the data collected in this thesis can be compared to previous research.
3. Both scales have an acceptable level of validity and reliability.
4. Both scales are free to use.

The decision to measure people's general level of altruistic behaviour was based on a review of the literature on motivations of the crowd in crowdsourcing contexts and the motivations of people participating in volunteering contexts (See chapter 2, sections 2.5 and 2.6). Both these literatures highlighted the importance of the altruism factor for people to participate in crowdsourcing and volunteering projects. It was therefore decided to measure students' general level of altruism using the SRA scale to investigate its effect on students' participation in the DescribeIT project.

1.3 Research Key Findings

The focus of this programme of research, as explained earlier, was to investigate students' motivation to participate in a socially responsible crowdsourcing project, the DescribeIT Project, to support blind and partially sighted students by describing images in digital learning resources. The series of empirical studies reported in this thesis contributes to our general understanding of what motivates people to participate in socially responsible crowdsourcing projects by investigating which intrinsic and extrinsic motivational factors could motivate students to participate in the DescribeIT project.

The key findings of this programme of research are the following:

1. The studies in this thesis have shown that it is very important to study actual behaviour as well as self-reported predictions of behaviour, as these can be very different. It is clear from the results of the studies that positive interest and willingness to participate to the DescribeIT project as found in the self-report phases was not matched with actual participation when the same participants were asked to describe the images in the second phase of the studies.

2. Overall, students rated the likelihood of their own participation to the DescribeIT project significantly higher than that of other students on their course.
3. The self-report phases showed that altruism and monetary rewards were the leading self-reported motivational factors for students in the UK. In addition to being interested in the accessibility area; the time expected from them to participate in the project being not too long; the ease of providing the descriptions; seeing how their participation would be helpful to blind or partially sighted students; participation increasing their understanding of the learning resources and knowing a blind or partially sighted person. However, actual behaviour showed that students are primarily motivated by monetary reward to participate in the DescribeIT project.
4. A multi-method approach to measure the quality of image descriptions including Signal Detection Theory was developed in this programme of research. The results showed an average level of accuracy (the degree of which the image description is precise in relation to the content of the image); low level of efficiency (the extent to which a description provides accurate information in the least number of words); and extremely low level of errors (reflects the frequency of false positives in an image description) of the image descriptions produced.
5. When students were paid they produced the highest number of descriptions per student. However, the quality of their descriptions was average and comparable to non-student participants (MTurkers and inidividuals participating in a citizen science platform).
6. The results suggest that there was no relationship between the number of images produced per participant and their attitudes toward people with disability (as measured by the Interaction with Disabled Persons Scale (IDP)) and their general altruistic attitudes (the Self-Report Altruism Scale (SRA)).

In addition to these findings, this thesis highlights additional implications for crowdsourcers to consider when creating projects for socially responsible tasks, as follows:

1. Make the task easier by breaking difficult and daunting tasks into smaller components. In this research asking participants to tag objects in an image

increased participation in comparison to asking them to describe the whole image.

2. If the task is complex or unfamiliar, training crowd members to do the task rather than just providing guidelines or instructions on how to complete the task may improve the quality of the work submitted. In this research training the students to do the image description task significantly improved the quality of the descriptions in comparison to only providing them with set of guidelines and an example.

1.4 Ethical Statement

All the studies conducted for this programme of research were guided by the principles of ethical research with humans. All the studies described in this thesis followed the ethical principles of 'Do No Harm', 'Confidentiality', and 'Informed Consent'. The studies were approved by the Physical Sciences Ethics Committee of the University of York.

Do No Harm: None of the participants in any of the studies conducted in this programme of research were put in any harmful situations. All the studies were designed in such way that participants would not be subject to any risk. Participants had the right to stop or withdraw their participation at any time. Participants were informed that they would not be penalised for withdrawing nor they will be questioned on why they withdraw.

Confidentiality: All the data collected were kept anonymous and confidential. In the results reported participants' anonymity was maintained at all times and a coding system was used to keep individual identities hidden when quoting participants' thoughts. Data was stored in password protected systems to protect them from unauthorised access.

Informed Consent: All participants were informed about the aims, procedure and tasks of the specific study they were invited to undertake. In Studies 1 to 10, which were all conducted online, participants were informed of the appropriate information about aims, anonymity and confidentiality before the study began and were allowed to ask any questions they might have by emailing the researcher. Study 3 Phase 3 was conducted face-to-face. For this study, participants were briefed about its aims and their rights before the interview started, and they were requested to read and sign an informed consent form. All participants were debriefed after the interview session was completed. Similarly, in Study 10 participants were provided with a copy of the study information document, consent form, and the project information page (see Appendix F.1.4 and F.1.5). Participants were briefed about the importance of the DescribeIT project, the study's aims, procedure and tasks. At the end of the study participants were debriefed and thanked for their participation.

1.5 Thesis Structure

The thesis is organized in the following chapters:

Chapter Two presents the Literature Review.

Chapter Three presents the DescribeIT project.

Chapter Four presents Study 1, the exploratory study which investigated what university students thought would motivate them to participate in a crowdsourcing project to support blind and partially sighted students.

Chapter Five presents Studies 2, 3, and 4, which investigated how different motivational factors can affect students' participation. These studies also compared what students said would motivate them with their actual behaviour.

Chapter Six presents Study 5 which investigated the of motivations and behaviour of students from a different culture to participate in the socially responsible crowdsourcing project.

Chapter Seven presents Study 6 which investigated the influence of monetary rewards on students' participation rate, and Study 7 which investigate the quality of the image descriptions produced by Amazon Mechanical Turk workers in comparison to those produced by students in Study 6.

Chapter Eight presents Studies 8 and 9 which investigated the effect of simplifying the image description task to an image tagging task on participation rate.

Chapter Nine presents Study 10 which investigated the influence of face-to-face training on the quality of image descriptions. It also investigated the effects of quality control messages in the instructions for the image description task. It also investigated students' perceptions of the difficulty of the image description task and whether this was altered by the training.

Lastly, Chapter Ten presents the overall discussion and conclusions of this programme of research.

Chapter 2
Literature Review

2.1 Introduction

This chapter presents the most relevant research and theory in the field of crowdsourcing and motivation. The chapter is divided into four parts. The first part is an introduction to the concept of crowdsourcing, followed by an overview of early crowdsourcing initiatives to support people with disabilities. The second part presents theories of motivation. The third part presents the measurements used to assess motivation and techniques used in research to manipulate motivation. The last part is an overview of empirical studies on the motivation of crowd members to participate in crowdsourcing projects.

2.2 Crowdsourcing

Crowdsourcing is a newly developed concept first introduced by Howe (2006). Howe defined crowdsourcing as *the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call* (Howe, 2006).

Although Howe was the first to define crowdsourcing, there are many other crowdsourcing definitions that emphasise crowdsourcing in different ways. According to Estelle-Arolas (2012), it is the adaptability of crowdsourcing that makes it a powerful tool. Estelle-Arolas (2012) examined more than 40 definitions of crowdsourcing in 209 documents, including conference papers, workshops, journal articles, books, and technical reports. From these he developed a general definition that can be applied to all types of crowdsourcing:

> *Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the*

crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken (Estelles-Arolas and Gonzalez-Ladron-de- Guevara (2012), 2012, p9)

To understand the mechanism of crowdsourcing, it is best to introduce its terminologies with an example.

2.2.1 Terminology in crowdsourcing

Crowdsourcing generally includes three main components: The crowdsourcer, the platform and the crowd members or crowd.

The term crowdsourcer, refers to the individual or organization initiating the crowdsourcing process, seeking the help of people in a task or group of tasks through a crowdsourcing platform. The platform in crowdsourcing is often a web-based organization which facilitates crowdsourcing services to mediate communication between the crowdsourcer and the people doing the tasks. The crowd members or crowd are the group of people who are recruited through the platform to undertake the crowdsourcer's task or group of tasks. As Estelle-Arolas (2012) noted, crowdsourcing covers a wide range of tasks and hence different incentive mechanisms are applied to motivate the crowd members to undertake tasks. For example, the platform Amazon Mechanical Turk[1] (MTurk), an online marketplace which enables crowdsourcers (known as requesters on MTurk) to post tasks (known as Human Intelligence Tasks or HITs on MTurk) for crowd members (known as MTurk workers, MTurkers or simply workers on MTurk) to undertake in exchange for micropayments (small amounts of money).

2.3 Crowdsourcing to Support People with Disabilities

Long before the invention of computer-based crowdsourcing, able-bodied people were giving their time and skills to support people with disabilities. For example, sighted people have long read to blind people, and more recently this became organized in talking newspapers and audiobook services. With the advent of the internet and the World Wide Web, crowdsourcing has facilitated many easier ways for able-bodied people to support people with disabilities and numerous interesting innovations have been proposed. Early initiatives in this area were surveyed by Bigham, Ladner and Borodin (2011).

[1] https://www.mturk.com/mturk/welcome

One of the most interesting early initiatives was from Von Ahn and Dabbish (2004) who developed a system whereby sighted people could provide descriptive tags of images on the web which would be very useful for blind and partially sighted web users. Strictly speaking, this might not be a crowdsourcing project, as the task was presented as a fun game, but it had the characteristics of using the power of the crowd to solve a problem for particular groups of web users. In a very similar approach Von Ahn et al. (2006) developed the Pethch game, which allows people to add image captions as result of playing the game. Takagi et al. (2008) took a more classic crowdsourcing approach to the problem of the accessibility of websites, addressing not just the issue of image tagging, but all the problems that disabled users encounter in using the web. Using their Social Accessibility Project, disabled web users could report a problem on a website. With a simple shortcut key command, they could open a dialogue box to report the problem, which would then be posted for a member of the crowd to solve.

In an educational context, Synote, a web application for annotating and captioning multimedia by Wald (2010) allows students to create synchronised bookmarks which contain their notes and tags to an audio or video recording.

Another group of crowdsourcing projects have addressed the problems that people with mobility and visual disabilities have in navigating in the physical environment. Hara, Le and Froehlich (2012) created a crowdsourcing project to tag Google Street View images with information about the accessibility of sidewalks. Cardonha et al. (2013) used crowdsourcing for a similar aim of creating accessibility maps of cities. In a second project Hara et al. (2013) used crowdsourcing to collect information about bus stops for blind people.

Finally, one of the most ambitious crowdsourcing projects to date to support disabled people is VizWiz, developed by Bigham et al. (2010). This project aims to provide nearly real-time answers to questions about visual information for blind people, such as the labels on food packets, dials on appliances and the colours of objects. This is achieved by the blind person taking a photo of the information with their mobile phone which is then relayed to MTurk workers in real time using a special service, quikTurkit.

These initiatives have shown that crowdsourcing has the potential to be a useful means of undertaking socially responsible activities to support people with disabilities. However, many technical and practical challenges need to be addressed to achieve a successful crowdsourcing project in this area. One of those challenges is the need for an appropriate crowd of people to undertake the work

Research on crowdsourcing projects to support people with disabilities has not investigated what motivates people to participate in these projects. To the best of the author knowledge, only one of the projects discussed above investigated the crowd's motivations. Takagi et al (2008) interviewed participants in a pilot scheme of their Social Accessibility Project about their motivations, although the number of interviews is not mentioned. Only brief qualitative results were reported, with participants mentioning the ranking of crowd members on the portal page and the appreciative comments from disabled users as effective motivators.

2.4 Motivation Theory, Research, and Practice

Before moving on to the research on the motivation of people to participate in different crowdsourcing projects, it is important to provide a definition of motivation and discuss how to assess people's motivation.

2.4.1 Definition of Motivation

Motivation has been for decades an important subject for economists, sociologists and psychologists. It has been examined as a psychological act involving desired goals, state or outcomes. The word "motivate" is derivative from the Latin word "movere" which means "to move". The psychologist Geen (1995) defines motivation as *the initiation, direction, intensity and persistence of human behaviour*. Motivation according to Pritchard and Ashwood (2008) *represents how we allocate our energy to satisfaction our needs* (p7). In more general terms, motivation can be seen as a description of why a person does something. In the everyday world, people become a part of something to meet their own needs, whether these needs are seeking a career, networking with other professional people, improving ones' skills, the desire of being a part of something big, experience, personal satisfaction, obligation, curiosity, entertainment, or social recognition.

2.4.2 Motivation Theory

Many theories have been proposed in the field of motivation to explain why and how people's behaviour is activated. The literature on motivation theories is far too extensive to be summarized in this section, thus I will only highlight some of the most commonly theories used to understand motivation in the context of crowdsourcing.

One of the most popular and accepted theory of motivation is Self-Determination Theory (SDT). According to SDT (Deci & Ryan, 1985) there are different types of motivation that influence people's behaviour. In SDT Deci and Ryan proposed different motivations arranged on a continuum of three levels from the highest to lowest levels of self-determination, as follows:

1. Intrinsic Motivation: refers to engaging in a task because of it is inherently pleasurable and satisfying (i.e. a person would participate for the sake of the task itself).
2. Extrinsic Motivation: refers to engaging in a task because of its outcomes (i.e. the task is used as a means to lead to an outcome), SDT proposes different types of extrinsic motivation, including:
   - Integrated Regulation: this type of extrinsic motivation is the most self-determined, but still not considered an intrinsic motivation because they are done not for their inherent enjoyment. It occurs when identified regulations been evaluated and brought to one's other values and needs.
   - Regulation through Identification: this type of extrinsic motivation is more self-determined, it occurs due to valuing the behaviour goals or regulation.
   - Introjection Regulation: *involves taking in a regulation but not fully accepting it as one's own* (Ryan and Deci, 2000, p72). Usually the behaviour is performed to avoid guilt or to attain pride.
   - External Regulation: the behaviour is preformed because of a possible reward or punishment.
3. Amotivation: people who are amotivated lack motivation to engage in any activity. Amotivated persons have little or no intention of undertaking a task, which results from not desiring the outcome, or not valuing the task, or not feeling competent to undertake the task (Ryan and Deci, 2000).

In addition, SDT is based on the fundamental idea that people strive to satisfy three fundamental psychological needs: *autonomy*, *competency* and *relatedness* (Deci and Ryan, 2012; Ryan, 2012). The need for *autonomy* refers to self-endorsement or self-determination. The need of autonomy is a sense of willingness when undertaking a task. In other words, autonomy is the desire to experience self-regulation (Ryan, 2012). The need for *competence* refers to a confirmation of one's self-esteem and mastery of experience. *Competence* in SDT is the challenging feeling when undertaking the task.

Lastly, the need for *relatedness* refers to the environment during the execution of a task. *Relatedness* means the need to be connected, close to other people and groups, and experience their care and trust.

To propose a model for motivation in the context of crowdsourcing Kaufmann et al. (2011) adapted different theories including Deci and Ryan's SDT theory (Deci & Ryan 1985, 2000), work and education motivation theories (Hackman & Oldham 1980, Weis 1995) and an open source software development model (Lakhani & Wolf, 2005). Their model of motivation (See Figure 2.1) is composed of two main motivational factors: *intrinsic* and *extrinsic* motivation, which are then broken down into five motivation categories. Very similar to SDT, *Intrinsic Motivation* refers to the inherent satisfaction of doing the task and *Extrinsic Motivation* refers to the desired outcome received upon completing the task.

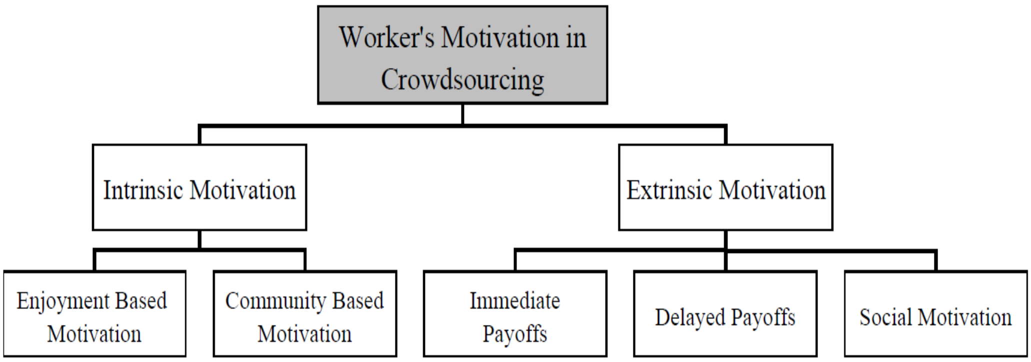


Figure 2.1 A model of motivation in crowdsourcing (source: Kaufmann et al., 2011)

As illustrated in Table 2.1, within the Intrinsic Motivation group there are two categories: Enjoyment Based and Community Based Motivation. The Extrinsic Motivation group includes three categories: Immediate Payoffs, Delayed Payoffs and Social Motivation.

Another interesting model to discuss in relation to motivation and behaviour is Fogg's (2009) Behaviour Change Model which is based on three principal factors: motivations, abilities, and triggers. Fogg's model proposes that to perform a specific behaviour a person must be sufficiently motivated, must have the ability to perform the behaviour, and be triggered. Fogg defines a *trigger* as something that tells people to perform a target behaviour now. He categorises *triggers* into three types: spark, facilitator, and signal. A *spark trigger* can be used when people are not motivated to do the behaviour, this can be in a form of a text or video. A *facilitator trigger* can be used with people who are motivated but lack ability to perform the behaviour, to help make the behaviour easier for them to do, it also can be in a form of a text or video. A *signal trigger* is suitable for people who are motivated and have the ability to perform the target behaviour but need a reminder. Fogg argues that the three factors (motivation, ability, and triggers) must occur at the same time in order for the person to perform the behaviour.

Table 2.1 Examples of Intrinsic and Extrinsic Motivations (source: Kaufmann et al. 2011)

	Category	Constructs	Source
Intrinsic motivations	Enjoyment based motivation	*Skill Variety*: The diversity of skills that are needed for solving a task. *Task identity*: Refers to the extent a worker perceives the completeness of the task she has to complete. *Autonomy*: the degree of freedom that is allowed to the worker during task execution. *Direct feedback from the work*: the extent to which a sense of achievement can be perceived during or after task execution. *Pastime*: acting just to "kill time/avoid boredom".	(Hackman and Oldham 1980, Brabham 2008; Ipeirotis 2010)
	Community based motivation	*Community identification*: the act of workers guided by the subconscious adoption of norms and values from the crowdsourcing platform community which is caused by personal identification process. *Social contact*: motivation caused by the sheer existence of a community that offers the possibility to foster social contact.	(Lakhani & Wolf 2005, Lindenberg 2001, Brabham 2008, 2010)
Extrinsic Motivations	Immediate payoffs	Payment: motivation because of the monetary remuneration received for completing a task.	(Lakhani & Wolf 2005)
	Delayed payoffs	Delayed payoffs are benefits which can be applied strategically to generate future material advantages and include: *Signalling*: Refers to the usage of actions as strategic signals to others. *Human capital advancement*: The possibility to develop skills that could be useful to generate future material advantages.	(Lakhani & Wolf 2005; Weiss 1995)
	Social Motivation	*Action significance by external values*: captures the significance of an action concerning the compliance with values from outside the Crowdsourcing community that is perceived by the worker when contributing to the contributing to the community or working on a task. *Action significance by external obligations and norms*: motivation induced by a third party from outside the platform community that traces back to obligations a worker has or social norms he or she wants to comply with in order to avoid sanctions. *Indirect feedback from the task*: motivation caused by the prospect of feedback about the delivered working results by other individuals.	(Deci & Ryan 1985, Ryan & Deci 2000, Hackman and Oldham 1980)

2.4.3 Measuring Motivation

This section will present some of the methods used to assess and measure motivation. In addition, it will highlight some of the challenges faced by researchers in measuring motivation.

Motivation cannot be observed or directly recorded (Toure-Tillery and Fishbach 2014), thus measuring motivation was always a challenging undertaking. Researchers as Toure-Tillery and Fishbach, (2014) said *measure motivation in terms of observable cognitive (e.g., recall, perception), affective (e.g., subjective experience), behavioral (e.g., performance), and physiological (e.g., brain activation) responses and using self-reports.* (Toure-Tillery and Fishbach, 2014, p328)

2.4.3.1 Self-Report measures of motivations

Probably the simplest approach to understand people's motivation is to ask them directly what motivates them to do a specific action. McDonald (2008) noted the popularity of obtaining self-reported data, she said it *makes complete sense – if I want to learn more about somebody, why would I not go directly to that person?* (p76). She examined the advantages and disadvantages of using a self-report method. For example, the advantage of using this method is that it is relatively inexpensive and quick to collect data from a large number of participants. For example, Deci and Ryan have developed many questionnaires to collect self-reports of different aspect of motivation[2], these are available free of charge for researchers and each take only a few minutes for participants to complete.

In spite of these advantages of using self-reports measures of motivation, researchers should use this with caution, especially if this was the only measurement used to assess motivation. Self-report may not reflect people's true motivations, for either intentional or unintentional reasons. Reporting one's motivation requires an understanding of one's own psychological state (Toure-Tillery and Fishbach 2014), so a participant might unintentionally report inaccurate motivations. In addition, participants usually tend to respond in ways that makes them look good to the researcher or in way that they think it is what the researcher wants, this is known as experimental bias (Rosenthal and Rosnow, 2009). In this case participants might intentionally under-report behaviours which might be viewed as inappropriate by researchers, or over-report behaviours which can be viewed as appropriate. Another concern is social desirability bias (Rynes, Gerhart, and Minette 2004), this bias describes participants' tendency to answer question in a manner that is socially accepted. Fisher (1993) noted that social desirability bias has been found in all self-report measures across nearly all social science research.

Furthermore, in the self-report method researchers must take a great care when structuring a questionnaire, as Schwarz (1999) stated that in self-report *minor changes in question wording, question format, or question context can result in major changes in the obtained results* (p. 93).

[2] http://selfdeterminationtheory.org/questionnaires/

2.4.3.2 Behavioural Measures of Motivation

Researchers often use people's behaviour as an indicator of their motivation. For example, Toure-Tillery and Fishbach (2014) noted that research on motivation often measures people's choices, the duration of preforming specific tasks, or the type of performance on tasks. The following section will discuss each of these behavioural measures of motivation.

The term *choice* is used by Toure-Tillery and Fishbach (2014) to describe *the act of selecting between objects (e.g. apple versus cookie) and courses of action (e.g. donating, exercising)* (p333). For example, Fishbach and Zhang (2008) examined how the presentation of items, (e.g. healthy food versus unhealthy food; academic versus leisure activities) could influence the dynamic of self-regulation. The results showed that when the items are presented in one image and next to each other, participants tended to choose the unhealthy food or the leisure activities (i.e. tempting items). Whereas, when participants were presented with the items in set of two images apart from each other, participants tended to choose the healthy or academic items (i.e. goal items).

Speed as Toure-Tillery and Fishbach (2014) explained, refers to the duration that an individual requires to complete a particular task. Overall time and speed can be useful measure of motivation. The time required to complete a task, the overall time invested in contributions to a crowdsourcing project over a short or long period of time could help researchers in understanding participants' behaviour in crowdsourcing projects (e.g. Mao, Kamar and Horvitz 2013; Goncalves et al. 2013; Difallah et al .2015). Mao, Kamar and Horvitz (2013), investigated participants' engagement with online crowdsourcing tasks, to develop statistical models which could predict the level of engagement with the task for upcoming participants. The authors used a set of data collected from Galaxy Zoo, a citizen science project, in which participants classify images of galaxies. Participants on Galaxy Zoo complete many tasks over a period of time and these tasks are carried out in one or more sessions, divided by breaks. Consequently, the authors have used the time spent on each task and the break time to understand how the participants perceive tasks and to predict when participants are going to stop participating in Galaxy Zoo.

Goncalves et al. (2013) investigated people's motivation to participate to a non-paid crowdsourcing service on public displays with eight different motivational factors. Participants were asked to count malaria-infected blood cells on a set of 30 images. The authors used behavioural measures to analyse participants' motivation, including time spent on the task. The results showed a significant relationship between completion time and the motivation.

The level of performance is another measure of behavioural motivation, which includes the accuracy of the work produced, how much of the work has been done, or the persistence with which an individual is doing a task (Toure-Tillery and Fishbach 2014). Many researches have investigated the performance of their participants in order to understand the effectiveness of their designs or their manipulations of motivational factors. The following example illustrate the use of performance measures to understand participants' behaviour.

Thaler, Simperl, and Wölger (2012) evaluated two conditions (Amazon MTurk vs. gamification) to investigate which was better in term of quality of output and costs (in the sense of development time and money). The authors used a categorization task to build a schema of Wikipedia articles. In the first condition (Amazon MTurk) participants were paid USD 0.74 per answer, whereas in the second condition participants were engaged in a game to do the task for free. In term of the quality, the data collected by both conditions were high in quality, meaning participants in both conditions executed similar level of task performance. However, in term of cost, the development time for the gamification condition was higher, and the authors pointed out to the challenge of applying appropriate player-retention strategy (to make sure players return to play). By contrast, in the Amazon MTurk condition the cost the development time was low, but there was finical expense for undertaking the task.

Given these different ways of measuring motivation and the complexity of the concept, it is important to use a mix of measures to capture various aspects of motivation (Toure-Tillery and Fishbach 2014). The validity of data collected by relying only on one method has been questioned, Donaldson and Grant-Vallone (2002) cited examples of the extensive research sharing this concern. Not only does relying on one method to measure motivation threaten the validity of the research as Antin and Shaw (2012) shown, but also it will hinder the development of online behaviour theories.

2.4.4 Techniques for manipulating motivation

This section outlines the commonly used techniques to implement motivational factors in empirical studies in crowdsourcing settings. Lee et al. (2013) surveyed some of these motivational techniques.

2.4.4.1 Textual Instructions to manipulate motivation

Many researchers used textual instruction to manipulate motivational factors in their studies (e.g. Chandler and Kapelner 2013; Rogstadius et al. 2011; Shaw, Horton, and Chen 2011; Goncalves et al., 2015). The textual instructions given by the researchers to the participants before participation is designed to influence their participation. For example, Chandler and Kapelner (2013) asked 271 MTurk workers to label medical images, participants were assigned to one of two conditions: The *meaningful* condition, in which participants were told that their participation will help researchers label tumour cells, or the *zero-context* condition, in which participants were told that they will be labelling "objects of interest", without telling them the purpose of their task. The results showed that providing a meaningful context does not necessarily increase the quantity or the quality of image labels.

Rogstadius et al. (2011) asked 158 MTurk workers to identify blood cells infected with malaria parasites. The authors hypothesised that an intrinsic motivational factor, such as helping others, would increase the quality of the work produced, and an extrinsic motivational factor would not. The study was 2x3 design. The motivation manipulation were two levels of instructions: The intrinsic motivation instructions emphasized that the task was being done for a *non-profit* organization, and the extrinsic motivation instructions emphasized that the task is done *for-profit* organization. With three levels of monetary rewards: 0, 3, and 10 cents USD. The work submitted was assessed by measuring the speed of completion and accuracy. Results showed that participants in the intrinsic motivation condition produced work of a higher quality than those of the extrinsic motivational condition. In addition, paying participants more did not improve their work quality. However, there was no significant differences in the completion time between the conditions.

2.4.4.2 Gamification to manipulate motivation

Deterding, Khaled, Nacke, and Dixon, 2011) defined the term of *gamification* as *the use of game design elements in non-game contexts* (p10). Many researchers, implement gamification aspects as motivational factors in crowdsourcing contexts (Morschheuser, Hamari, and Koivisto, 2016), such as, rewarding participants with points (e.g. Dumitrache et al. 2013), badges (e.g. Bowser et al. 2013; Feyisetan et al. 2015), stars, or virtual currency, implementing level of difficulties (e.g. Dumitrache et al. 2013; Feyisetan et al. 2015), progress bars (e.g. Chandler and Horton, 2011), and leader boards (e.g. von Ahn and Dabbish, 2008; Bowser et al. 2013). In their review

Morschheuser, Hamari, and Koivisto (2016) investigated how different gamification implementations can increase motivation and participation in crowdsourcing. They noted that when crowdsourcers outsource tedious and repetitious tasks, they tend to use simple gamification implementations, whereas, when they outsource creative tasks they tend to use a richer set of gamification implementations.

2.4.4.3 Feedback to manipulate motivation

Giving feedback to participants is considered a form of reward. However, some researchers (e.g. Lee et al. 2013) categorize feedback as form of gamification, I do agree with them that for example a progress bar could be considered an immediate form of feedback, but it only highlights the quantity of the work done, not its quality. For this reason, I will discuss feedback in the sense of providing quality check on the work done rather its quantity.

Dow et al. (2012) investigated if feedback improve the quality of work produced by participants in crowdsourcing platform, by testing *Shepherding*, a feedback system for crowdsource work. 207 MTurk workers participated in the study, workers were assigned to one of three conditions: No feedback, self-feedback, and expert feedback. Workers were asked to write a consumer review for six products they use. In the No feedback condition, workers advance form writing a review to the next one, with no feedback on their reviews. In the self-feedback condition, workers can assess their review before submitting. In this condition workers assess their work by grading rubric. In the expert-feedback condition, workers receive feedback from experts on their consumer review, before submitting the next one. The expert used the same grading rubric used in the self-feedback condition. The results showed that both the self-feedback and expert-feedback conditions yielded better overall work than the no feedback condition.

2.5 The Motivation of the Crowd

In the context of crowdsourcing, understanding what motivate people to participate in different projects receiving a great deal of attention. Although researchers have not investigated the motivation of the crowd in crowdsourcing projects to support people with disabilities, motivational factors that lead people to participate in crowdsourcing projects in general have been investigated. These can be divided into intrinsic and extrinsic motivations. In the context of crowdsourcing, extrinsic motivation means that the crowd members are motivated by factors external to the crowdsourcing task; these might be financial rewards, social recognition, or job opportunities. Whereas intrinsic motivation means that crowd members are motivated by factors related to the task itself, such as the satisfaction they got from undertaking the task.

Lakhani et al. (2007) explored the motivation of the crowd to participate in the scientific problem-solving project InnoCentive. In an online survey 357 crowd members rated 16 motivational factors. The results showed that while InnoCentive crowd members were motivated by monetary rewards, they were significantly more motivated by intrinsic motivation factors, such as the joy of the intellectual challenge of problem solving. Interestingly, both the extrinsic motivation of money and intrinsic motivations were significantly correlated with being a successful problem solver. In addition, having free time to participate in InnoCentive was significantly and positively correlated with success. However, other extrinsic motivational factors such as social motivations or the desire to beat others to solving the problem were negatively correlated with problem solving.

Oreg and Nov (2008) explored the motivation of the crowd in the domain of open source software (Sourceforge) and content (Wikipedia) developments. In an online survey with 185 crowd members, they found that members of the open source software group rated self-development and reputation building higher that members of the open source content group. However, members of the open source content group rated altruism higher than the open source software group.

Brabham (2008, 2010) examined the motivation of the crowd in artistic design projects in a series of studies. In his 2008 study he investigated the motivation of crowd members on *iStockPhoto*, an online royalty-free photography agency for amateur photographers. An online survey with 651 “iStockers” investigated their motivations for participating, and showed that the desire to make money, improve one’s creative skills, and to have fun were the strongest motivators, whereas passing the time and building a network of friends were the weakest motivators for participation at iStockphoto.

In a second study, Brabham (2010) conducted instant messaging interviews with 17 members of *Threadless*, an online t-shirt company that crowdsources the design process of its product through an ongoing online competition. Brabham found that crowd members had a mix of intrinsic and extrinsic motivations, including money, developing creative skills, the potential to leverage freelance design work, and the love of the Threadless community. Most interestingly, one motivation was addiction to the

Threadless community, whereby the members of the crowd at Threadless see themselves as part of Threadless rather than just external contributors.

Hossain (2012) targeted 400 crowdsourcing platforms to identify the key activities and incentive mechanisms used. He found that only 27.6% of the platforms used intrinsic motivations, whereas 72.3% are using extrinsic motivations to motivate the crowd members. Of these, approximately 50% of the platforms use monetary rewards as the extrinsic motivation. These results show that intrinsic motivation is much less common on crowdsourcing platforms than extrinsic motivation, and that monetary rewards are the most common extrinsic motivation.

There are two weaknesses in the existing research that must be addressed in studies which wish to contribute to answering the research question proposed in this programme of research. The first weakness is that most studies conducted to investigate crowd members' motivations rely on online survey methodologies in which respondents self-report their motivations. This methodology according to Antin and Shaw (2012) is vulnerable to social desirability bias, as studies' participants may respond to survey questions in way that they believe reflect social norm and is seen as socially desirable and appropriate. As a result, we must also study the behaviours of people while performing crowdsourcing tasks and compare these with what they say has motivated them to perform the tasks.

Secondly, most of the existing motivation studies examine the motivation of the crowd members in relation to a particular task. While there may be overlaps between what would motivate people to participate in crowdsourcing projects in general, crowd members in different projects doing different tasks may prioritise the importance of different motivational factors differently. In the case of the studies that are proposed for this programme of research, what motivates people to participate in an existing commercial crowdsourcing task will not necessarily motivate them to participate in a socially responsible crowdsourcing project to support disabled students. Hence, it is important to examine the motivation of the crowd undertaking socially responsible related crowdsourcing tasks in these studies.

2.6 The Motivation of the Volunteers

As mentioned in the previous section, researchers have not investigated the motivation of the crowd to participate in socially responsible crowdsourcing projects. Hence, the motivation of the crowd to participate in crowdsourcing projects in general was discussed instead. However, understanding what motivates people to participate in volunteering in socially responsible projects in contexts other than crowdsourcing may help in understanding people's motivation to participate in socially responsible projects in the crowdsourcing context. Accordingly, the following section will outline research on people's motivation to participate in socially responsible projects in non-crowdsourcing contexts.

The functional approach to volunteering is a theoretical framework to understand people's motivation to participate in volunteering. It was developed by Clary, Snyder and their colleagues (See Clary, Snyder, and Stukas, 1996 for detailed review) and then have used it for many years. This approach as Clary and Snyder (1991) note is "*concerned with the reasons and purposes that underlie and generate psychological phenomena - the personal and social needs, plans, goals, and functions being served by people's beliefs and their action's* (Clary and Snyder 1991, p123). The approach proposes that people volunteer to fulfil psychological functions. Clary, Snyder, and Stukas (1996) identified and validated six categories of motivations (or psychological functions) that can be served by volunteering:

1. Values function: in which people participate in volunteering to express or act on values important to themselves.
2. Understanding function: whereby people participate in volunteering to increase their knowledge and develop their own skills.
3. Enhancement function: in which people participate in volunteering to engage in psychological development and enhance their self-esteem.
4. Career function: whereby people participate in volunteering to gain experiences that will benefit their careers.
5. Social function: in which people participate in volunteering to be part of a group and get along with social groups they value.
6. Protective function: other people participate in volunteering to cope with inner anxieties and conflicts.

As Houle, Sagarin, and Kaplan (2005) noted, the functional approach is a *multi-motivational perspective*. This means that an individual can be motivated by more than one function, and a group of volunteers can be motivated by different function to do the same activity. In addition, the same individual can be motivated by different functions when participating in different volunteering activities.

Clary, Snyder, and Stukas (1996) analysed data collected in the United States by the "National Survey of Giving and Volunteering" in light of the functional approach. They used the Volunteer Functions Inventory (VFI), an instrument used to assess the functions served by volunteering (Clary et al. 1998). 2671 Americans who had participated in volunteering activities in the previous year were asked to identify the area/s in which they had volunteered from a list of 15 areas (e.g. health, education, religious, and human services). The aim was to investigate the relation between motivation and volunteering in different activities. The results showed that the motivational functions measured by the VFI significantly related to the type of volunteering activities they chose to undertake. The results particularly highlighted the importance of the value motivation to participate in volunteering in general, and that other motivations then guide participants to choose which specific volunteering activity they undertake.

Bussell and Forbes (2002) reviewed the literature on volunteering to understand the key motivational factors that can influence recruiting and retaining volunteers. They noted that altruism and wanting to help others is often the most frequent motivational factor given for volunteering. However, there is a debate in the literature about whether altruistic motivation is the only driving motive for volunteering. Interestingly, the review highlighted very similar motivational factors to the ones reported in the crowdsourcing context, including passing the time, improving one skills, advancing in one's career, enhancing job opportunities, and gaining academic credits. Other motivational factors that could potentially influence people's decision to volunteer were being asked by someone they value, if they have a family or friend involved with the organization, and for religious beliefs. Less common motivational factors for volunteering included *wanting to wear a uniform, perks obtained, mixing with celebrities, health and fitness, travel opportunities* (Bussell and Forbes, 2002, p11).

Houle, Sagarin, and Kaplan (2005) investigated whether certain volunteering tasks such as data entry, typing letters, and reading to a blind person differentially satisfy certain motivational functions, and whether participants prefer tasks that are expected to satisfy their personally relevant motives. 112 psychology students completed the VFI, then they were presented with descriptions of eight volunteering tasks and were asked to rank them from the most preferred one to the least preferred. Afterwards participants were presented with descriptions of six volunteering motivational factors along the descriptions of the eight tasks. They were asked to evaluate the extent to which each task would satisfy each of the six volunteering motivational functions. The results suggested that participants could differentiate tasks based on the motives they satisfy. In addition, the results showed that participants preferred tasks that matched their own most important motives.

2.7 Personality Traits

People decision to participate in a socially responsible crowdsourcing project may be driven by intrinsic or extrinsic motivational factors or mix of both. However, their decision can be also influenced by individual differences and personality traits. This means that people's likelihood of participation to a crowdsourcing project aimed to support disabled students could be influenced by numerous different dimensions of personality traits. The most widely accepted model of classifying personality traits is the Big Five which is also known as Five Factor model: Openness, Conscientiousness, Agreeableness, Extraversion, and Neuroticism (McCrae and John, 1992; John and Srivastava, 1999).

Although investigating the effects of personality traits on participation in the DescribeIT Project was out of the scope of this programme of research, it worth highlighting the five dimensions of personality traits which might influence participation in socially responsible crowdsourcing projects.

Big Five Personality Traits:

Openness: openness is short for "openness to experience". This personality trait includes how much someone is open or resistant to change, and how inclined they are to cultural norms. People with high scores on this dimension tend to be imaginative, independent, and interested in variety. In contrast, people with low scores tend to be practical, conforming, and interested in routine.

Conscientiousness: conscientiousness refers to someone's degree of organization and self-discipline. People with high scores on this dimension tend to be self-disciplined, dutiful, organized, and careful. In contrast, people with low scores on this dimension tend to be disorganized, careless, and impulsive.

Extraversion: this personality trait includes how social someone is, and the degree of warmth and love someone shares with others. People with high scores on this dimension tend to be sociable, fun-loving, and affectionate. In contrast, people with low scores on this dimension tend to be retiring, sombre, and reserved.

Agreeableness: includes how kind someone tends to be or how sympathetic to others they are. People with high scores on this dimension tend to be compassionate, cooperative, trusting, and helpful. In contrast, people with low scores on this dimension tend to be self-interested, antagonistic, and uncooperative.

Neuroticism: this trait relates to how nervous someone tends to be or the degree of someone's self-confidence. People with high scores on this dimension tend to be anxious, insecure, and emotional unstable. In contrast, people with low scores on this dimension tend to be calm, secure, and self-satisfied.

Faullant et al. (2016) investigated the influence of personality traits on people's decision to participate in a crowdsourcing competition. 57 participants participated in the competition and completed an online questionnaire to assess their personality traits. In addition, 112 participants participated only in the online questionnaire to assess their personality traits but did not participate to the crowdsourcing competition (although it is not clear that these two groups of participants are actually comparable). Participants in the crowdsourcing competition scored significantly higher in openness, extraversion, dispositional trust, and trait competitiveness than those who did not participate in the competition.

2.8 Scales for Measuring Motivation and Attitudes

A number of scales were used throughout the studies in this programme of research, the following sections will provide an overview of the scales used.

It was hypothesised that people who are considerate of others and altruistic are more likely to participate in a socially responsible crowdsourcing project to support blind and partially sighted students than those who are less considerate and altruistic. Therefore, students' (and others) sense of altruism was measured using the Self-Report Altruism Scale (SRA) (Rushton et al. 1981). The SRA 20-item scale in which respondents rate the frequency with which they have engaged in altruistic behaviours on 5 point Likert items. Scores range from 20 (least altruistic) to 100 (most altruistic). Rushton, Chrisjohn and Fekken (1981) had shown that SRA scale has acceptable reliability and validity. Some of the items on the SRA were considered not appropriate for use with student (and other) samples in the UK, Libya and Saudi Arabia (e.g. I have helped push a stranger's car out of the snow), so four items from the SRA were dropped, making a 16-item scale.

It was also hypothesised that students' (and others) attitudes towards people with disabilities could affect their behaviour in a crowdsourcing project to support disabled students. Therefore, it was important to measure students' (and others) attitudes in this regard. This was done with the Interaction with Disabled Persons Scale (IDP) (Gething 1994). The IDP is a 20-item scale to measure participants' perception of their interactions with people with disabilities on 6-point Likert items (1= agree very much to 6 = disagree very much). Maclean and Gannon (1995) proposed that the IDP consists of two sub-scales "Discomfort" about people with disabilities and "Sympathy" toward people with disabilities. Acceptable levels of reliability and validity have been reported by Gething (1991) for this scale.

Lastly, to assess students' motivation when they are actually engaged with the image description task, the Situational Motivation Scale (SIMS) (Guay et al. 2000) was used. The term *situational motivation* refers to "the motivation individuals experience when they are currently engaging in an activity" (Guay et al. 2000, p176). The SIMS is a 16-item scale to assess why participants are engaged in an activity on 7 point Likert items (1=not at all, to 7= exactly). It comprises four sub-scales:

Intrinsic Motivation: level of engagement in the activity for its own sake.

Identified Regulation: extent to which the activity is perceived as being chosen by oneself.

External Regulation: extent to which the activity is perceived as being regulated by rewards or negative consequences.

Amotivation: perception that the activity lacks consequences, positive or negative.

Chapter 3

A Platform for Developing Socially Responsible Crowdsourcing Project

3.1 Introduction

In this Chapter, the platform that was used in the empirical studies of crowdsourcing in this programme of research will be described.

A wide variety of platforms were investigated for undertaking the empirical studies in programme of research. The main goal was to set up a general environment that could be easily modified as needed to support a variety of studies. The other key requirements for the platform were:

1. The platform must allow for voluntary contributions from crowd members, that is not require a micro-payment system. The key reason for this is that compulsory monetary compensation for crowdsourcing tasks would be a confound in any study looking at other motivational factors.
2. Easy to configure task settings such as: adding and removing tasks for a study, or changing task priority.
3. Easy to capture data about participant demographics and task behaviour (e.g. number of task per participant, number of task per day).

The Crowdcrafting.org platform was the only platform investigated which captured the full range of requirements, while remaining easy to modify with minimal web programming. Furthermore, Crowdcrafting already has a large collection of volunteers who could be useful to invite to participate in some of the studies in this programme of research.

A project called DescribeIT was developed that delivers images to crowd members who then create descriptions of images suitable for blind and partially sighted people.

3.2 DescribeIT: A Crowdsourcing Project to Describe Images to Support Blind and Partially Sighted Students

DescribeIT is a crowdsourcing project to support blind and partially sighted students in higher education by the description of images in digital learning resources. The vision is that lecturers would be able to upload to the project their digital learning resources that require descriptions of the images before giving them to a class of students. Once the images are described by sighted students, the lecturer would check the quality of the descriptions, and if desired, edit the descriptions, possibly selecting and/or editing the best description, if a number are provided for the same image. In addition, lecturers could give feedback to the students on their descriptions. The feedback provided could improve students' skills for future descriptions and could potentially increase students' understanding of the teaching materials. The materials could then be provided in advance of lectures and seminars to blind and partially sighted students, as they often find it very helpful to study materials in advance of these sessions. Of course, the materials, with their image descriptions, would also be available during and after sessions, for interactive use and revision.

To help sighted students create good image descriptions, the DescribeIT project information page provides guidelines on how to describe images for blind and partially sighted people and an example description of a typical image. The guidelines used in this project were developed from those developed by Chen (2013) for describing images on museum websites for blind and partially sighted people (See section 3.3.1 for more details).

The DescribeIT Project was also designed to incorporate motivational factors based on Deci and Ryan's self-determination theory of motivation (see Chapter 2, section 2.4.2), the factors being autonomy, competence and relatedness. To support autonomy (students' sense of willingness when undertaking the image description task), students have the choice of creating an account on the Crowdcrafting platform or to login anonymously. In both cases, it is possible to collect basic information about the students, such as username code or IP address. In addition, students are free to describe as many (up to the maximum number of images available) or as few images as they wish, at any time they wish. Once students decide to start describing images, they are presented with a digital learning resource such as a PowerPoint slide and a text-box in which to type their description of the image on the slide (See Figure 3.1). Students are able to create a description and submit it and then go to the next image for description. If they do not wish to describe a particular image (e.g. if they do not understand it), they can skip the image, and go to the next image. The skip function was also implemented to increase students' sense of autonomy.

Figure 3.1 Add image description page for the DescribeIT project

Students' competence is supported by implementing a progress bar to show each individual contribution (how many image descriptions have been created and how many images are still available to describe). This feature was later removed as the result of a pilot study (Study 2, see 5.3.2.5) in which participants thought that it might hinder participation when there are lots of images to describe. Unfortunately, the overall number of tasks available in the project (i.e. number of slides) and the total number tasks done (i.e. images described) by all participants together is presented by the Crowdcrafting platform. This feature cannot be removed by a project developer, so could not be removed for this research.

Relatedness, or people needs of being connected to others, is incorporated in the the DescribeIT project by facilitating human-to-human interaction (Zhang, 2007). Students are able to directly contact the crowdsourcer (the researcher) by sending a direct email when needed. In addition, students are able to leave comments about a specific image (anonymously or by their user name) or share their descriptions on social media.

Since the project was developed on an open source platform, anyone who can access the platform was able to access the project and participate. Thus, a password system was implemented to make sure that only the students targeted for any particular study can access the project and participate.

The DescribeIT project allows for the collection of the following pieces of information:

1. The image descriptions.
2. The number of images each student describes in a session, per day and in total.
3. The images that students choose not to describe.

In all the studies reported in this thesis the same settings presented here were used, except for the motivational instructions and the example image description. The manipulation of the motivational factors instructions were presented both in the recruitment email sent to students inviting them to participate in a study and also on the project information page. The example image description also was different in each study (as the example was a slide from the material used in the project). The guidelines on how to describe images for blind and partially sighted students, text font and size were identical across all the studies to remove any potential bias.

For Studies 8 and 9, another crowdsourcing project, TagIT, was created with a different task. Details about the TagIT project will be discussed in Chapter 8, section 8.2.2.4.

The digital learning resources used in the DescribeIT project varied from study to another, as in Studies 2 ,3, 4, and 10 students were asked to describe images from materials that were being used in their teaching at that time. In studies 5 ,6, 7, 8, and 9 the participants were not students or were students from a range of different departments, so teaching materials that do not require a specific knowledge to describe its images were used (further details about the digital learning resources used will be presented in the material section of each study). Thus, the number of images available to be described varied between studies.

3.3 Preliminary Studies

After developing the DescribeIT project, two preliminary studies were carried out. The first preliminary study developed guidelines on how to describe images for blind and partially sighted students. The second preliminary study evaluated the usability of the DecribeIT project.

3.3.1 Preliminary Study A: Guidelines on how to describe images for blind and partially sighted students.

3.3.1.1 Introduction

To help students create good image descriptions for blind and partially sighted students, it was decided to adapt an existing set of guidelines from the literature developed by Chen (2013), for describing images on museum websites for blind and partially sighted people. Chen's guidelines were based on semi-structured interviews with 17 visually impaired people, in which she asked what they would like to know about an image and what they thought should be included in an image description. Chen had a first version of the guidelines evaluated by expert and non-experts, who also were asked to create image descriptions using the guidelines. The evaluation showed that more examples were needed. The quality of the descriptions produced were assessed through interviews with eight visually impaired people, by asking them to rank three different descriptions (short and long) of four images. The results suggested that the short descriptions (on average 34 words length) should provide the most interesting features of the image, whereas, the long descriptions (on average 132 words length) should provide more details.

However, the guidelines developed by Chen (2013) were to describe images on museum websites. Thus, it was essential to adapt these guidelines for digital learning materials and evaluate whether the adapted set of guidelines were suitable to describe images from such materials for blind and partially sighted students.

3.3.1.2 Method

Based on Chen's (2013) guidelines the researcher wrote a first version of guidelines for the describing images in digital learning resources for blind and partially sighted students. These were ordered from the general to the specific and were broken down into numbered and bulleted lists in an active voice. The first draft then was assessed by Professor Petrie to make sure that the set of guidelines were appropriate to the DescribeIT project. In addition, she reviewed the guidelines to make sure they were written in clear, concise, simple language, and that they are applied to all types of images used in digital learning materials.

The final draft of the guidelines was assessed by five participants, including three men, and two women. All participants were researchers working in the accessibility area and three of them have been working very closely with blind and partially sighted people.

Participants were asked to:

- Read the guidelines, and underline what they thought was important to consider when describing images in digital learning materials to blind and partially sighted students.
- Review an image description example.
- Describe an image according to the guidelines and the example.
- Assess the appropriateness of the guidelines to the DescribeIT project in open ended question.

3.3.1.3 Results and Discussion

As a result of the assessment of the first version of the guidelines made by Professor Petrie, the language used was improved and the guidelines were made shorter and more to the point. An example of both a short and long image description were constructed.

The final draft assessment showed that participants highlighted all the essential elements to include in an image description to a blind or partially sighted person, including: objects, people, locations, purpose of the image; colours, and what is happening in the image. Participants found the guidelines clear and succinct. They agreed that the use of language was appropriate for a formal document. The examples included provided a check of understanding in what should be included in a useful image description.

The final version of the guidelines can be found in Appendix A.1

3.3.2 Preliminary Study B: Usability of the DescribeIT Project

3.3.2.1 Introduction

The study was particularly motivated by the study by Brabham (2012), in which he examined the motivation of people participating in the Next Stop Design, a crowdsourcing application for transit planning. He conducted interviews with 23 participants, and found that the good usability and low barriers to entry were motivational factors that lead people to participate to the application. As discussed earlier in this chapter, the barriers to entry to the DescribeIT project were minimised as much as possible (free to use, with or without registration), but I also wished to investigate the usability of the DescribeIT project. A design with poor usability could be a barrier in any study examining the motivation of people using the DescribeIT project. Thus, it was decided to evaluate the usability of the DescribeIT project, to make sure that the usability of the project is satisfactory.

3.3.2.2 Method

To get a quick sense of the usability of the DescribeIT project the System Usability Scale (SUS) (Brooke, 1996) was used. Participants of the study were six postgraduate students at the University of York, comprising three men and three women.

Participants were asked to use the DescribeIT project to describe images and then were asked to complete the SUS in relation to the project.

3.3.2.3 Results and Discussion

The mean SUS score was 87.5 (SD= 6.1). According to Sauro (2011), a score of over 80.3 is in the top 10% of usability scores, so this means that participants found the DescribeIT project highly usable.

In conclusion, the design of the DescribeIT project as the study has shown was perceived by participants as usable design.

Chapter 4

Students' Perception of their Motivations to Participate in a Socially Responsible Crowdsourcing Project

4.1 Introduction

The first study (Study 1) in this programme of research investigated what university students think would motivate them to participate in a crowdsourcing project to support visually disabled students.

Study 1 was an exploratory study because crowdsourcing is an emerging phenomenon that is not yet well investigated. As discussed in Chapter 2, section 2.5 there is a lack of research into what motivates people to participate in socially responsible crowdsourcing projects. Most of the existing motivation research discussed in Chapter 2, section 2.5 examines the motivation of the crowd members in relation to a particular task. While there is an overlap between what would motivate people to participate in crowdsourcing projects in general, the crowd members in different projects doing different tasks will be driven by different motivations. In the case of the studies that are proposed for this programme of research, what motivates people to participate in an existing crowdsourcing task will not necessarily motivate them to participate in a socially responsible crowdsourcing project for disabled students. For this reason, it was important to investigate what students think would motivate them to participate to this particular task, of describing images for blind and partially sighted students using the DescribeIT project.

Study 1 used a self-report technique, an online questionnaire, which is a popular research methodology in a variety of disciplines, including HCI (Lazar, Fing, and Hochhesiser, 2017). Like any research methodology, an online questionnaire has several advantages and disadvantages. For this particular study, having the questionnaire online meant that students could respond to the questionnaire at any time suited them, this flexibility could increase participation rate in comparison to other self-report methods. However, one threat to the validity of any research employing a self-report method is the social desirability bias, in which respondents tend to respond in the way that they believe would be desirable to others. It is concerning how frequently self-report measures are used in crowdsourcing motivation research (See Chapter 2, section 2.5), particularly given the very little research that has been done to mitigate or assess this bias in research on motivations in participation in crowdsourcing. Subsequent studies in this programme will address this issue in detail. In this study, it will be addressed by triangulating answers to different questions about motivations.

4.2 Method

4.2.1 Design

The study investigated how university students self-report different motivational factors that would affect their participation in a crowdsourcing project to describe images for their blind and partially sighted peers, the DescribeIT project. An online questionnaire was used, comprising the following sections: an explanation of the DescribeIT project; questions about their perception of the likelihood of their own participation in the project and that of other students; open-ended questions about what factors would motivate them and other students to participate in the project; ratings of to what extent a set of known motivational factors for crowdsourcing would motivate them to participate in the DescribeIT project; and demographic and background information.

The rationale for the Study 1 was as follows:

1. To establish a baseline of self-reports of motivations to compare with actual behaviour later in the thesis.
2. To investigate what the students think would motivate them to select motivators to manipulate in subsequent studies.
3. To triangulate what students thought would motivate them (in the open-ended questions) and their ratings of known motivations for crowdsourcing.

4.2.2 Respondents

A total of 1640 students in departments in a range of disciplines at the University of York received an email inviting them to take part in the study. 271 students responded, giving a response rate of 16.5%. The respondents comprised 98 women and 172 men (one student preferred not to identify their gender). Their age range was from 18 to 51 years old, with a mean age of 20.9 years (SD = 5.1). The University of York is a very international institution, and while majority of participants were from the United Kingdom (178, 65.7%), the rest were from other countries. This included 6.3% (17) from China, 5 (1.8%) from each of Malaysia, the USA, and India; and 4 (1.5%) from each of Poland, Bulgaria and Spain. The remaining participants (18.0%) were spread relatively evenly across 25 other countries: Bangladesh, Canada, Chile, Cyprus, Egypt, France, Germany, Greece, Ireland, Italy, Jamaica, Jordan, Kenya, Latvia, Libya, Malta, Mexico, Norway, Pakistan, Russia, Saudi Arabia, South Africa, Sweden, the Netherlands, and Zimbabwe.

All respondents reported using social media, such as Facebook. However, only 17.0% (47) reported participating in crowdsourcing projects. These included Wikipedia, Amazon Mechanical Turk, Galaxy Zoo and many other projects. Respondents reported participating in these projects for different reasons (See Table 4.1) these include: monetary rewards, being interested in the topic, contributing to large projects, and the fun of the activity.

As an incentive to complete the questionnaire, respondents were entered into a prize draw for one of 10 Amazon vouchers worth £10 each.

Table 4.1 Students' comments on what motivated them to participate in crowdsourcing project in the past

Motivational Factor	**Students' comment**
Monetary rewards.	*I got paid for my skills.* (P18)
Being interested in the topic.	*It was on a topic I was interested in, and it was an opportunity to contribute to research in that field in a small way.* (P11)
Contributing to large project.	*Felt like it was a big project worth being part of* (P81)
The fun of the activity.	*I find it fun and interesting to contribute to these websites. It also lets you converse with some interesting individuals.* (P135)

4.2.3 Online Questionnaire

The online questionnaire consisted of four sections:

Introduction: was designed to give general information about the DescribeIT project, including the aim of the project, the importance of the project, an idea of how the project pages function, the task to be undertaken, and the type of images to be described.

Likelihood and motivation to participate in the DescribeIT project: consisted of four questions, which investigated students' perceptions of their likelihood of participating in the DescribeIT project, on 7 point Likert item (1= Not at all likely to participate, and 7= Very likely to participate) and their own statements of their motivations to do so. It also asked them to estimate the likelihood that other students on their course would participate, on 7 point Likert item (1= Not at all likely to participate, and 7= Very likely to participate) and what they think would motivate other students on their course to do so. The actual wording of the questions was:

Q1: How likely is it that you would participate in this project? [7 point Likert item]

Q2: What would motivate you to participate in this project? [Open-ended question]

Q3: How likely is it that other students on your course would participate in this project? [7 point Likert item]

Q4: What do you think would motivate other students on your course to participate in this project? [Open-ended question]

Motivational factors: a set of 12 motivational factors was presented (see Table 4.2), taken from the research on the motivation of the crowd in crowdsourcing projects. Students rated on 7 point Likert items (1= not at all, 7= very much) how much each factor would motivate them to participate in the image description project. For each factor, students were also asked to explain in an open-ended question why they had given that rating. At the end of this section students were asked to include any last thoughts about what would motivate them. The questions in this section were as follows:

Q5: Please answer the questions below about whether these factors would motivate you to participate in the image description project

Example: Motivational factor 1:

Your sense of altruism, wanting to help other students [7 point Likert item]

Why did you give that particular rating? [Open-ended question]

Q6: Do you think there are any other factors which would motivate you to participate in the image description project? If so, please describe briefly.

Demographic and online activities questions: this section collected demographic data about respondents' age, gender, department of study, the highest degree they hold and country of nationality. This section also collected data about the use of social media and previous experience with crowdsourcing.

The full set of questions can be found in Appendix B.1.3

Table 4.2 A list of the 12 motivational factors used in this study.

Motivational factor	**Source**
Your sense of altruism, wanting to help other students (Altruism)	Hossain (2012), Nov (2007)
Improving your academic skills (Academic Skills)	Brabham (2008, 2010, 2012), Hossain (2012)
Being paid for your efforts (Money)	Brabham (2008, 2010), Hossain (2012), Ipeirotis (2010), Lakhani et al. (2007)
Enhancing your job opportunities in the future (Job Opportunities)	Brabham (2010, 2012), Hossain (2012)
Being connected with other students on your course (Being connected)	Brabham (2008), Hossain (2012)
The fun and entertainment of the activity (Fun/Entertainment)	Brabham (2008, 2012), Hossain (2012), Nov (2007), Ipeirotis (2010),
To pass the time (Pass time)	Brabham (2008), Hossain (2012), Ipeirotis (2010)
Knowing that you are contributing to a large project (Contributing)	Brabham (2012),
The social recognition you would receive (Social recognition)	Brabham (2012), Hossain (2012)
Drawing attention to your skills (Attention)	Brabham (2008)
Being in a competition with other students (Competition)	Morschheuse, Hamari, & Koivisto (2016)
Getting academic credits (Academic Credits)	Wald (2010)[3]

4.2.4 Procedure

A recruitment email was sent out (see Appendix B.1.2) and a reminder email was sent out five days later. A further round of recruitment was conducted several months later,

[3] Wald (2010) did not investigate the effect of the motivational factor, getting academic credits, however, he suggested the use of this motivational factor to motivate students to participate in SyNote crowdsourcing system.

to increase the sample size. On each occasion, the online questionnaire was available for one week.

4.2.5 Data analysis

Not all the variables in the data collected were normally distributed, so it was decided to use nonparametric tests for all analyses.

Qualitative analyses of the two open-ended questions about students' motivations to take part in the DescribeIT and the open-ended question on their concluding thoughts were carried out using content analysis (Hsieh and Shannon, 2005). Any text which referred to distinct motivational factors was tagged by codes representing discrete motivational factors. This was carried out by two independent researchers, the researchers then worked together until they reached an agreement on the motivational factors categories.

4.3 Results

Students rated their likelihood to participate in the image description project and the likelihood that other students in their course would participate in the project, both on a 7 point Likert item (1=Not at all, 7=Very likely). Students gave a rating mean of 4.8 (SD=1.5) that they would participate in the project. However, students thought that likelihood of their peers to participate in the project less than them, with rating mean of 4.1 (SD=1.7). Figure 4.1 shows the pattern of mean ratings for these two questions. To investigate if the likelihood of students' ratings of the likelihood of participation was significantly above the midpoint of the rating scale, a one-sample Wilcoxon signed rank test was conducted comparing the ratings with the midpoint rating of 4 (See Table 4.3). This showed that students' ratings of their own likelihood to participate in the pr oject were significantly higher than the neutral midpoint (W=7.3, $p < .001$). However, their ratings for their estimation of the likelihood that other students would participate were not significantly different from the neutral midpoint of the rating scale. A Wilcoxon signed-rank test showed that, participants rated the likelihood of themselves participating in the project statistically significant than from the likelihood of their peers participating in the project (Z = -5.3, $p < 0.000$).

Students were also asked to state what they thought would motivate them and what they think would motivate other students in their course to participate in the project in open-ended answers. The content analysis revealed the categories of motivations that would motivate students to participate in the project and what they thought would motivate other students. These are summarized in Table 4.4. The most frequently mentioned categories were financial rewards, wanting to help others, for some students it was important to know that there are actual visually disabled peers who could benefit from the project, being interested in accessibility, participations do not require long time, the ease of providing the descriptions, knowing a blind or partially sighted person, improving skills and understanding of the teaching materials.

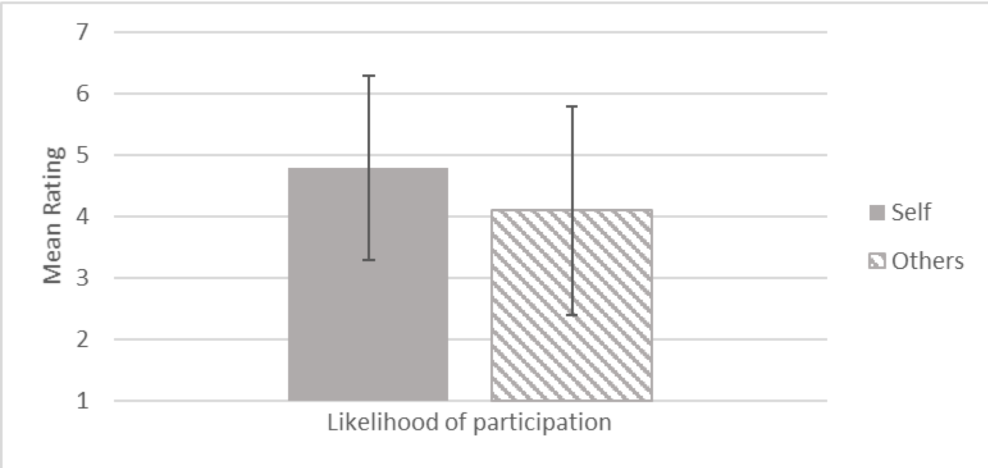


Figure 4.1 Mean ratings and stander deviation of likelihood of participating in a crowdsourcing project to describe images, for self and other students

Table 4.3 Mean ratings, median ratings, stander deviation and the results of one-sample Wilcoxon signed rank test of students' rating of their own and others likelihood of participation.

Likelihood of participation	**Mean**	**Median**	**SD**	**W**	***p***
Self	4.8	5.0	1.5	7.3	0.000
Others	4.1	4.0	1.7	1.2	n.s.

The most frequently mentioned motivational factor was getting some sort of financial rewards for their participation, where 92 (33.9%) students thought that would motivate them, a total of 144 (53.1%) students thought that would motivate other students. Some students who commented on the importance of this factor said:

I would be motivated by money vouchers. (P8)

Financial benefit (P17)

Table 4.4 Categorization of factors which students think would motivate themselves and other students to participate in the crowdsourcing project (from open-ended question)

Motivational factor	**Self** N (%)	**Others** N (%)
Financial Rewards	92 (33.9%)	144 (53.1%)
Wanting to help peers (explicitly mentioned knowing there is an actual blind or partially sighted peer)	81 (29.9%) (6, 2.2%)	32 (11.8%) (1, 0.4%)
Interested in the accessibility area	22 (8.1%)	4 (1.5%)
Does not require long time to participate	20 (7.4%)	9 (3.3%)
Easy to provide descriptions	18 (6.6%)	12 (4.4%)
Knowing a friend or a family member who is blind or partially sighted	15 (5.5%)	5 (1.8%)
Improving ones' skills and understanding of the teaching materials	10 (3.7%)	6 (2.2%)
Getting feedback about their contributions	9 (3.3%)	6 (2.2%)
Improving CV	8 (3.0%)	3 (1.1%)
Increase participants' awareness about the topic	7 (2.6%)	18 (6.6%)
Asked by their lecturers to take part in the project	6 (2.2%)	3 (1.1%)
Having fun	6 (2.2%)	2 (0.7%)
The images are related to their courses	4 (1.5%)	3 (1.1%)
Receiving reminder emails to participate in the project	3 (1.1%)	0 (0.0%)
Being in a competition with other students	3 (1.1%)	3 (1.1%)
Getting academic credits	2 (0.7%)	5 (1.8%)
Having disability themselves	2 (0.7%)	0 (0.0%)
Social recognition	1 (0.4%)	3 (1.1%)
Leader board	1 (0.4%)	1 (0.4%)
To pass the time	1 (0.4%)	0 (0.0%)

The second most frequently mentioned motivational factor was wanting to help disabled students. A total of 81 (29.9%) students thought that wanting to help others would motivate them, in comparison to only 32 (11.8%) students who thought this factor would motivate other students. Interestingly, six students (2.2%) explicitly mentioned that their motivation to help others will be only activated if they knew there is a blind or partially sighted colleague, in compare to only one student who thought this might motivate other students. Students said:

the satisfaction of helping somebody else get a better learning experience (P6)

Knowing I was actually helping partially sighted/blind people, no point otherwise. (P55)

if they were more aware of other student being partially sighted, for example if they could see what their effort was being put into (P75)

Being interested in the accessibility area was thought by 22 (8.1%) students to be a factor thought could motivate them to participate to the DescribeIT project, whereas, only four students (1.5%) thought that could motivate other students. Students commented:

I would probably participate as it is useful and important that resources be as accessible as possible. (P2)

Improving accessibility (P101)

Aspects of the properties of the image description task were thought by students to influence their participation to the DescribeIT project, this included, that the task does not require long time to complete (Self=20, 7.4%; Others= 9. 3.3%) and the ease of providing the descriptions (Self=18, 6.6%; Others= 12, 4.4%). Students thought that it is more important to themselves in compare to other students to do the task in short time, and they also though it is more important for other students to be able to provide descriptions easily in compare to themselves.

I'd like to say that just making things accessible to my peers would motivate me but I'm just so busy with my own work. So if this task didn't take longer than 30 minutes per week (5 mins every day) (P112)

The easiness of describing the images (P97)

Interestingly, a student suggested simplifying the task by asking questions about the image instead of asking to describe the whole image, he said:

I am unmotivated to participate by the open endedness [endless] of the text box. I feel that with a box that big I wouldn't knwo [know] what to write. I would be more motivated *had the text box been split into the categories of what the object is and what it is doing for example.* (P30)

Knowing a friend or a family member who is blind or partially sighted was thought by 15 (5.5%) students to be a motivational for themselves to participate, whereas, only five (1.8%) students thought that would motivate other students. Students said:

perhaps a friend or family member struggling with visual impairment. (P22)

Knowing a friend who was blind or partially sighted. (P26)

Ten students (3.7%) thought that improving their academic skills and understanding of the teaching materials would motivate them to participate to the DescribeIT project, in comparison to six (2.2%) students who thought that would motivate other students to participate.

If it would help me understand part of my course/studies better (P23)

A chance to learn something new. (P85)

Other motivational factors were mentioned as well, including:

Getting feedback about their contributions, a student said:

Feedback about how helpful my descriptions have been (P48)

Improving CV, a student commented:

The ability to put it on my CV as volunteering work. (P105)

Increasing participants' awareness about the topic, one student said:

Awareness of the project and a better understanding as to how it can help people who have visual impairment. (P28)

Asked by their lecturers to take part in the project, as a student commented:

If a person of authority (ex. a tutor) asks me to do it. (P84)

Having fun, one student said:

Make it fun and interesting. (P85)

The images are related to their courses, one student said:

Only describing images related to my course (P123)

Being in a competition with other students, one student commented:

being entered into a competition (P124)

Getting academic credits, a student commented:

Probably extra credits. (P112)

Social recognition, a student said:

Some recognition that the help was given (P104)

Leader board, a student said:

System with leader boards (P103)

To pass the time, a student said:

To fill in spare time (P126)

Students were then asked to rate to what extent each of 12 motivational factors would motivate them to participate, using 7 point Likert items (1= Not at all to 7=Very much). The results are illustrated in Figure 4.2 and Table 4.5. Overall, students rated their sense of altruism and wanting to earn money as the highest rated motivational factors to participate in the image description project. Whereas, to pass time and social recognition were the lowest rated motivational factors.

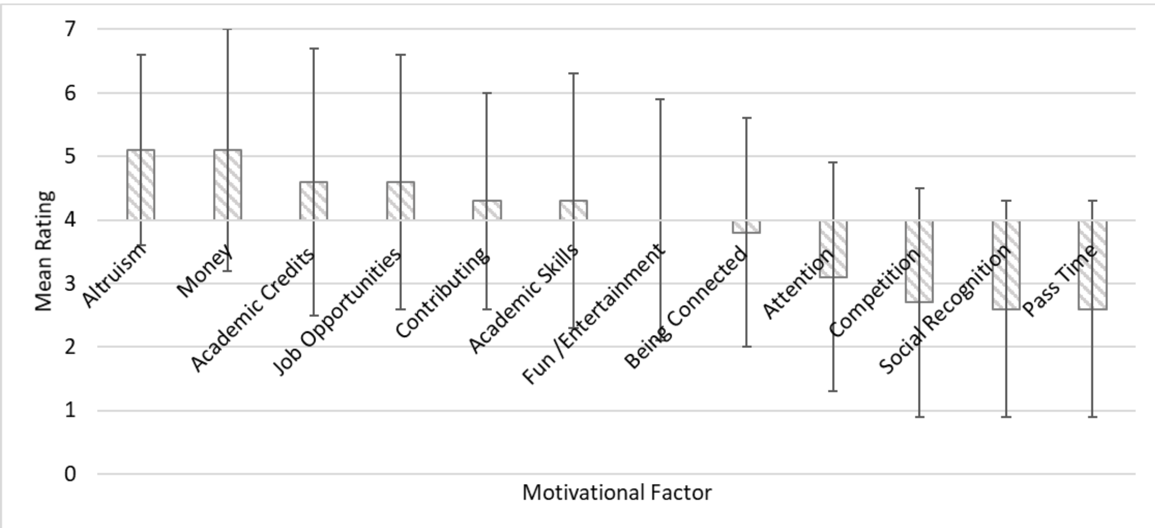


Figure 4.2 Mean ratings and stander deviation for the 12 motivational factors for self in relation to the neutral point on the rating scale

A series of one-sample Wilcoxon signed rank tests were carried out to investigate whether students' ratings of the 12 motivations are statistically different from the neutral midpoint rating of 4 (See Table 4.5). The results indicated that students rated altruism, money, academic credits, job opportunities, and contributing significantly above the midpoint. On the other hand, being connected, attention, competition, social recognition, and to pass time were rated significantly below the midpoint.

The following sections present more detailed results for each factor, with examples of students' comments about each factor.

Table 4.5 Mean, standard deviation and median ratings for the 12 motivational factors with results of one-sample Wilcoxon signed rank tests

Motivational Factor	**Mean**	**SD**	**Median**	**W**	***p***
Altruism	5.1	1.5	5.0	9.5	0.000
Money	5.1	1.9	5.0	7.8	0.000
Academic Credits	4.6	2.1	5.0	4.2	0.00
Job Opportunities	4.6	2.0	5.0	4.3	0.000
Contributing	4.3	1.7	5.0	2.8	0.004
Academic Skills	4.3	2.0	4.0	1.9	n.s.
Fun /Entertainment	4.0	1.9	4.0	-0.3	n.s.
Being Connected	3.8	1.8	4.0	-2.2	0.027
Attention	3.1	1.8	3.0	-7.4	0.000
Competition	2.7	1.8	2.0	-9.7	0.000
Social Recognition	2.6	1.7	2.0	-10.4	0.000
Pass Time	2.6	1.7	2.0	-10.3	0.000

Your sense of altruism, wanting to help other students

Students thought they would be significantly motivated by the altruism factor. Many students emphasized the thrill and joy of being able to help their peers. Students thought that by helping others there would be a good chance that others may help them back when they are in need. Also, students mentioned that helping students in need makes them happy about themselves. In addition, several students felt that their sense of altruism comes from their belief of equal education for all. Table 4.6 presents a range of comments from students at different rating levels.

Table 4.6 Comments from students about sense of altruism and wanting to help others

Ratings	Students' Comments
2	*To be honest, I'm probably too lazy to take part for this reason alone.* (P11)
3	*It's nice to help people but it wouldn't be effective use of my own time* (P17)
4	*I like helping people but I'm not very good at explaining.* (P35)
5	*I do care for how well other students are coping and how easily it is for them to study however I do have concerns about my study and so that would take priority.* (P14)
6	*Helping other students is emotionally rewarding and can be a great thing to do if you're in a similar situation to them* (P33)
7	*"I would happily contribute to making some other student's life easier. Especially if the contribution is as small as describing an image."(*P264)

Table 4.7 Comments from students about being paid for their efforts

Ratings	Students' Comments
1	*My time is more important to me than the amount that could reasonably be paid for completing this task.* (P245)
2	*I do not believe that you need to be paid for this project as I think it will be seen by most as a worthwhile use of there time.* (P21)
3	*As much as it would be nice to be paid, it takes away from the aspect of helping out fellow students and towards a paid job as such.* (P67)
4	*Being paid would be an added bonus, but wouldn't determine whether i participate or not.* (P100)
5	*it would depend on how much as well as the amount of effort needed to do the task* (P49)
6	*It would be an straightforward and relatively easy way to earn some money, which would be a good incentive*. (P73)
7	*As a student, money is an important part of the decisions I make every day. Living cheaply means taking opportunities for earning money on the side. I feel as if most students would agree.* (P23)

Being paid for your efforts

Students thought that they would be significantly motivated by the monetary reward factor. The following Table 4.7 presents some of the comments from students. The chance of making money through this project was a strong perceived motivation and seemed very appealing for the students. While many students acknowledged the importance of money for students and agreed on that monetary rewards would be a good reason for them to participate in the image description project, other students thought money would not motivate them, as their ultimate goal is to help others and make teaching materials more accessible.

Some of the students' comments regarding the monetary reward motivation showed that students thought about the task they would be doing (i.e. describing images) and evaluated the effort and time they though would be needed to undertake the task. This was reflected in their ratings. For example, two of the students who rated getting a monetary reward to be very motivating (i.e. giving it the highest rating of 7), explained their answers by saying *Because it is money for a very simple task* (P267) and *It may be a time-consuming task* (P128). Whereas, student P267 thought it is easy money to earn for a simple task, the other student (P128) thought the task might require more of his time, so an incentive should be offered to encourage people to take part. On the other hand, a student indicated that he is not at all motivated by a money reward (i.e. giving it the lowest rating of 1), due to lack of time. *I don't have time - it's not about money* (P169).

Getting academic credit

Students thought they would be significantly motivated by getting academic credit for participating in the image description project. Table 4.8 presents a range of comments from students on this factor. Interestingly, some students thought that getting academic credits would be good on their CVs, which might enhance their job opportunities in the future.

Enhancing your job opportunities in the future

Students thought they would be significantly motivated by the idea that they could enhance their job opportunities in the future by participating in the DescribeIT project. Examples of students' comments on this factor are presented in Table 4.9. Many students commented that participating in such a project would look good in their CVs especially if they were able to certify their participation, *If there was a way to certify this, so I could show employers, I would definitely be motivated* (P7). However, some students were not motivated by enhancing their job opportunities. The lack of interest in this motivation factor was for a number of reasons, such as lack of interest in this field, or because they thought that the skills that they would learn from this project will not help them to improve their job opportunities.

Table 4.8 Comments from students about getting academic credits

Ratings	Students' Comments
1	*"They would be superfluous to be needs."* (P239)
2	*Only a slight incentive as I feel getting a degree is much more important than an extra academic credit.* (P14)
3	*Would be great to receive credits for this but I still believe the main purpose of this activity is to help people* (P93)
4	*Would be nice, but doesn't determine whether I would participate or not* (P100)
5	*Any chance to get academic credits is likely to motivate most people* (P82)
6	*I would do it anyway, but gaining academic credits is quite motivating*. (P111)
7	*It may improve job prospects later* (P110)

Table 4.9 Comments from students about enhancing job opportunities

Ratings	Students' comments
1	*I prefer to enhance my opportunities building networks, improving my knowledge and skills. This will not be a motivator for me. (*P254)
2	*I don't think it would really enhance my job opportunities*. (P8)
3	*If it had the potential to enhance my job opportunities in a signnificant [significant] way then it would be something worth doing* (P62)
4	*It's always good to have something to add to your cv but I don't think this would contribute greatly* (P72)
5	*I would say that by improving my skills within the subject of image recognition would enhance my job opportunities in the future.* (P102)
6	*It would be very beneficial to be able to record such a project for my CV.* (P104)
7	*This would be something I may be interested in further on in my career when I would like to specialise.* (P146)

Knowing that you are contributing to a large project

Students thought they would be significantly motivated by the idea that they would be contributing to a large project. Comments from students in regard to knowing that they are contributing to a large project factor are presented in Table 4.10. Students who were motivated by this factor wrote about being part of something big to help others means that they are helping a large group of people. However, students who were less motivated by this factor thought that the impact of their participation would be less significant to a large project.

Table 4.10 Comments from students about knowing that they are contributing to a large project

Ratings	Students' comments
1	*For me this works the other way around. Since this is going to be a really big project I feel like my contribution will mean less. If it was much smaller I would be much more likely to participate.* (P8)
2	*I wouldn't use the end product or know about how the rest of the project was going.* (P17)
3	*I prefer being a large actor rather than a small contributor* (P25)
4	*I'd like to see such a project succeed* (P58)
5	*It will give a good sense of achievement when you have completed it.* (P18)
6	*Great to know that through participating in this activity, I am able to help make our world better and it would thus be considered an achievement not only for me but everyone else taking part.* (P93)
7	*Being part of something bigger is a nice feeling* (P7)

Improving your academic skills

Students were overall neutral about the motivation of improving their academic skills. Table 4.11 presents examples of comments from students about the improving their academic skills factor. Some students thought that describing the images in the learning resources could help them to understand them better. Some students were less motivated by this factor, and could not see how the image description activity would help them to improve their academic skills. Interestingly, a student noted a fear of providing poor descriptions to disabled students due to lack of necessary skills, which could negatively impact their learning.

Table 4.11 Comments from students about improving academic skills factor

Ratings	Students' comments
1	*Whilst one may revise learning material by participating, I don't think that would be even a major reason as to why they participate in the first place, as there are as many ways of revising as there are students* (P87)
2	*I don't see how my academic skills would be improved in a major way* (P62)
3	*I do not feel that it would improve my academic skills by any large margin, I also fear that if I was to get something wrong because my academic skills were lacking in an area I would be detrimenting other students learning* (P13)
4	*Chnace [chance] to learn something new* (P85)
5	*I think putting things into your own words will help your own understanding of a topic.* (P18)
6	*To describe an image, you'd have to understand what that image is showing, possibly re-inforcing your knowledge behind it.* (P67)
7	*I want to be able to do my job and do it well.* (P17)

Table 4.12 Comments from students about the fun and entertainment of the activity factor

Ratings	Students' comments
1	*I don't believe I would find the activity fun at all as I don't enjoy anything writing related.* (P15)
2	*I imagine it would get monotonous quite quickly* (P105)
3	*The activity of describing pictures may get repetitive over time, which would reduce the amount of fun and entertainment.* (P122)
4	*The activity itself would not be very fun as it is describing pictures, it wouldn't be a boring task but I wouldn't do it for the entertainment aspect.* (P104)
5	*It will be a challenge to put certain images into words and I enjoy problem solving.* (P18)
6	*It seems like an interactive enjoyable task as opposed to something meticulous and boring* (P36)
7	*A fun activity is intrinsically motivating.* (P26)

The fun and entertainment of the activity

Although students were overall neutral about the motivation of the potential fun and entertainment of the image description task, some students thought that the task would be fun to do because it could be a challenging task to them. While some students did positively comment that the image description task might be fun, other students were not convinced. For examples of students' comments see Table 4.12.

Being connected with other students on your course

Overall, students thought they would not be motivated by being connected with other students on their course. Table 4.13 presents examples of comments from students on this factor, the majority were not positive about it. However, some students commented positively about this.

Table 4.13 Comments from students about being connected with other students factor

Ratings	Students' comments
1	*I do not care about the other students on my course* (P14)
2	*I have no particular motivation to connect with others on my course. Instead I prefer interacting with people outside of my course.* (P22)
3	*I get that it's important, but I just don't really NEED it.* (P43)
4	*Not too interested in this, I don't need extra connections.* (P83)
5	*The project provides the environment of meeting other students on my course.* (P97)
6	*Because I find it a little bit difficult sometimes to connect with other students on my course.* (P107)
7	*I really like making things all-inclusive and accessible to everyone- I think having a strong bond with course mates is essential.* (P113)

Drawing attention to your skills

Overall students thought they would not be motivated by the idea of drawing attention to their skills, or indeed that the task needed skills which could be highlighted. However, some students saw some potential in this motivational factor. Examples of students' comments are presented in Table 4.14.

Table 4.14 comments from students about drawing attention to skills factor

Ratings	Students comments
1	*I think many or most people with sight will be able to describe images in the same way as myself, so I don't consider it a skill.* (P16)
2	*I feel like the task isn't extremely difficult and other task related to my course would draw attention to my skills better.* (P14)
3	*I don't really want to make it seem like I am better/worse than anyone else.* (P41)
4	*Descriptive skills are important but I don't think I need to draw much attention to them.* (P73)
5	*It could help boost my CV* (P58)
6	*It would both improve my skills and draw attention to my skills.* (P70)
7	*I like to know what I'm good at and what I need improving on.* (P72)

Table 4.15 Comments from students about being in a competition with other students factor

Ratings	Students' comments
1	*This project should be about helping more than competing, as the end goal is to assist those without sight. Not to win.* (P15)
2	*I am not a very competitive person.* (P22)
3	*I am not a fan of competition all that much.* (P28)
4	*I'm reasonably competitive*. (P32)
5	*Having a competition between other students would motivate me to contribute* (P62)
6	*If I did a better job than the other students and was recognized for it then that would motivate me quite a lot.* (P8)
7	*I believe the nature of the world is competitive, and a fair competition improves everyone involved.* (P26)

Being in a competition with other students

Overall, students thought they would not be motivated by the idea of being in a competition with other students. Although it was surprising to see that full-time university students are not feeling competitive, some students noted that they did not feel competition was appropriate in a project that was supposed to help their disabled peers. Nonetheless, some students did comment positively on this factor, some of these are presented in Table 4.15.

Table 4.16 Comments from students about the social recognition they would receive

Ratings	Students' comments
1	*This isn't something I would brag about. (P29)*
2	*Prefer to be anonymous* (P17)
3	*While it may be good to receive some social recognition, it's not something I would associate with that* (P103)
4	*I don't like the idea of doing something like this just for social recognition, though social recognition isn't a bad thing. (P73)*
5	*If I could be seen to be being very helpful by my friends, that would be great.* (P41)
6	Gain respect from others (P90)

Table 4.17 Comments from students about passing the time factor

Ratings	Students' comments
1	I have plenty of things, far more enjoyable than this, to do which will pass the time. (P14)
2	*I have way too much stuff to do all ready* (P20)
3	*I could probably find other ways to pass the time.* (P16)
4	*I would if I had a lot of time to kill, but I have other things I could procrastinate with.* (P24)
5	While there are better timewasters out ther[e] variety is the spice of life and I can imagine myself spending some time on this in-between other activities. (P30)
6	*Doing something useful with my time is important to me*. (P41)
7	*This could work well for me, if I needed a ten minute break from a project*. (P84)

The social recognition you would receive

Overall students thought they would not be motivated by the idea they would receive social recognition for participating in the project, mainly because they thought it was not appropriate to get recognition for a project that was supposed to help their disabled peers. Table 4.16 presents some of the comments on this factor.

To pass the time

Overall, students thought they would not be motivated by the idea of using the image description activity to pass the time. Table 4.17 presents some examples of comments from students on this factor. Students thought there are many ways to pass time in an enjoyable way in comparison to describing images. It is worth noting that students repeatedly mentioned in their comments that they often do not have spare time due to study requirements.

4.4 Discussion

The present study explored what students perceive would motivate them to participate in a crowdsourcing project to support their blind and partially sighted peers. Students were asked the likelihood that they would participate in such a crowdsourcing project and their perception of the likelihood of other students on their course participating. They were also asked what would motivate themselves and other students. Then they were asked to rate 12 motivational factors.

Overall, students showed an interest in participating in the socially responsible crowdsourcing project to describe images in electronic learning resources for their blind and partially sighted peers. Students rated their likelihood of participating as positive, significantly above the neutral point of the rating scale. However, students rated the likelihood of other students' participation to the project as neutral and significantly lower than their own likelihood. The fact that students rated the likelihood of their own participation significantly higher than that of other students on their course is interesting. This could be due to what social psychologists call the "fundamental attribution error" (Ross and Nisbett, 2011) which proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits. Thus, when asked to predict their own participation, students think of the specific situation of helping other students who are at a disadvantage, but when asked to predict the behaviour of other students, they think of the general helpfulness of other students, not in the context of helping disabled students. This was also noted when students commented on what would motivate them and what would motivate other students to participate in the project.

To mitigate the influence of social desirability bias, students were asked in open-ended questions about what would motivate them and what they think would motivate other students to participate to the DescribeIT project. A triangulation was carried out for the answers of these two questions. In addition, a triangulation of their answers about what would motivate themselves and their ratings of the same factors was carried out. This technique had yield some interesting results, as follows:

Firstly, when students were asked in the open-ended question what would motivate them and other students to participate, they noted a range of motivational factors, the two leading motivational factors were financial rewards and wanting to help others. The extrinsic motivational factor of a financial reward was the most frequently mentioned factor, 34% of the students thought that would motivate them in comparison to 53% who thought that would motivate other students. Whereas, the intrinsic motivational factor wanting to help others was the second most frequently factor, 30% thought that would motivate them in comparison to 12% who thought it would motivate other students. This showed that when students mentioned an extrinsic motivational factor they put more emphasis on its importance when they thought of what would motivate other students than in relation to themselves. In contrast, when they mentioned the intrinsic motivational factor they put more emphasized on its importance when they thought of what would motivate themselves in comparison to other students.

Secondly, the results of the present study showed that altruistic motivation was highly valued by students. This factor was the highest rated motivational factor and it was the second most frequently mentioned motivational factor in the open-ended question answered by students. The importance of this motivational factor was reported in previous studies (Kuznetsov 2006, Oreg and Nov 2008) which found people's sense of altruism is one of the leading factors for self-reported motivations for participating in socially responsible crowdsourcing projects such as Wikipedia.

Monetary reward was the second highest rated self-reported factor. It was also the most frequently mentioned factor in the open-ended questions. This finding agrees with the findings of Brabham (2008, 2010) and Lakhani et al. (2007) which both found that monetary reward was important as a self-reported motivation.

Other motivational factors, such as, improving one skills, enhancing CVs, the fun of the activity, being in competition, getting academic credits, and social recognition were thought by students to be less important than money and altruism, when they were asked what would motivate them to participate to the DescribeIT project in the open-ended question. However, when comparing the open-ended questions results to the ratings of the same motivational factors the results matched in some motivational factors and were contrary to others. These findings will be discussed in the following sections in relation to the findings of previous research.

The findings of the present study suggest that students' ratings were neutral about the fun and entertainment of the activity, which is contrary to some other studies (Brabham 2008, 2012, Hossain 2012, Kaufmann, et al. 2011). These studies found that the fun and entertainment was the leading self-reported intrinsic motivation factor in relation to their crowdsourcing tasks. This could be due to the nature of proposed task in the DescribeIT project, as students did not think describing an image as fun task to do, whereas Brabham's studies examined artistic design tasks.

The results about improving one's skills were not similar from those from Brabham's (2008, 2010, 2012) studies. The present study found that students' ratings on this factor were neutral. However, improving or developing artistic skills at iStockphoto (Brabham 2008) or Threadless (Brabham 2010) which he found were motivating factors on those project may not be exactly the same as improving students' academic skills, being more personal than career-oriented skills. Perhaps for this reason students did not think that this factor would motivate them to participate when they were asked in the open-ended question.

The motivation of getting academic credits has not been examined in any previous research. However, the present study showed a significantly positive self-report rating for this factor. Nonetheless, only two students thought that would motivate them to participate in the open-ended question.

While competing to solve difficult scientific problems might sound appealing to students, being in competition to help their blind and partially sighted peers was not. This was reflected both in the open-ended question and the students' ratings. Similarly, students did not think they would be motivated by the social recognition they would receive both in their ratings and their answers to the open-ended questions. This result contrasted with Brabham (2012) and Hossain (2012) results, which both reported the importance of social recognition.

Job opportunities, were thought to be significantly important by students when they were asked to rate this item. However, in the open-ended questions students did not explicitly mention job opportunities, but they did mention about being interested in the field, improving skills, and improving CVs which all can potentially enhance one's job opportunities. The opportunity to advance in one's career was reported to be an important motivational factor to participants of Next Stop Design as reported by Brabham (2012).

Previous research (Kuznetsov 2006, Brabham 2010) emphasized the importance of the community within the crowdsourcing projects and being connected to other participants in crowdsourcing projects. However, the results of the present study suggest that being connected to other students was not thought by students to be a important motivation to participate in the project. This particular finding is similar to Brabham's results from his study of iStockphoto (Brabham, 2008) in which he found a lack of connection between iStockphoto members due to trust issues. Similar to iStockphoto members (Brabham, 2008), students thought they would not be motivated by passing the time. As the respondents were full-time students they are often have much work to do and according to them if they have spare time they would spend it doing more enjoyable things. This was also reflected in their answers to the open-ended questions, in which students thought that they would be motivated to participate if their participation did not require dedicating a lot of time.

Students did not think they would be motivated by drawing attention to their skills, which contrasted with the results of Braham (2008) study, in which iStockphoto members reported the importance of drawing attention to their artistic skills. In addition, this comes at odds compared to the fact that students though they would be highly motivated by enhancing their job opportunities.

It is worth noting that students mentioned some motivational factors related to the nature of the DescribeIT project, such as being interested in the accessibility area. They also thought that knowing someone blind or partially sighted would motivate them to participate in the project. Obviously, that is a particular motivation when the crowdsourcing project is to help blind and partially sighted people, but it has wider implications for socially responsible crowdsourcing, that knowledge and contact with the target group could be very motivating and important.

Moreover, students thought that if describing the images is easily done and does not require too much of their time, that could motivate them to participate. The ease of use was also reported by Braham (2012), where he reported that participants of Next Stop Design were motivated by the ease of use of the system. Students' concerns about the time that their participation might take match their ratings of passing the time factor, which is understandable as explained earlier.

4.5 Conclusions

The results of this study suggest that students thought they would participate in the DescribeIT project, and that they are more likely to participate in the project than other students in their courses. This could be due to what social psychologists call the "fundamental attribution error". The first finding is encouraging for the future of the project to help blind and partially sighted students. However, the latter finding suggests the influence of social desirability bias, in the sense that students responded to the question about their own likelihood of participation in what they perceived as socially acceptable. In addition, the results present a first insight into what students think would motivate them to participate in a crowdsourcing project to support their blind and partially sighted peers. Students said they would be particularly motivated to participate in the crowdsourcing project by their sense of altruism and monetary rewards. In addition, students suggested new motivational factors that fall outside the 12 motivational factors examined, such as being interested in the accessibility area, the time expected from them to participate in the project, the ease of providing the descriptions, and knowing a blind or partially sighted person.

Thus, the results reinforce the notion that each crowdsourcing project has its own particular relevant motivations, although in line with previous research, this study also shows that there are self-reported motivations in common.

A key point to note is that the present study and most studies that have investigated crowd members' motivations have relied on respondents' self-report of their motivations. As Antin and Shaw (2012) note, this methodology is very vulnerable to the "social desirability bias", meaning participants may respond to questions in ways that they believe that they should, in ways that are socially acceptable. This was particularly noted in the current study when comparing students' open-ended questions answers to their ratings.

Chapter 5

Studies of a Socially Responsible Crowdsourcing Project with UK Students: Comparisons of Self-Reports of Motivation to Participate and Actual Participation

5.1 Introduction

Further studies in the programme of research with the same group of students who participated in Study 1 were planned to investigate students' actual behaviour in participating in the crowdsourcing project to provide image descriptions for electronic learning resources. In these studies, the findings of Study 1 would be compared with students' actual behaviour in the DescribeIT project, to examine the differences between their self-reported predictions of their motivations and their actual behaviour. However, unexpectedly, I was not able to get the teaching materials from the range of departments which were targeted in Study 1. For over a month my supervisor contacted lecturers and staff at different departments with little r esponse. Thus, I had to only carry out studies with departments that we have personal contacts with to get the teaching materials for the project.

King and Brunner (2000) noted that social desirability bias is more likely to occur when the topic under investigation is a socially sensitive one. Taking into account the social and normative implications when investigating the motivations of people to participate to a socially responsible crowdsourcing project to support blind and partially sighted students, one must not rely only on self-reported motivation to participate in such a project.

This chapter will present three studies which investigate UK students' actual participation in the DescribeIT project to provide image descriptions for digital learning resources. The three studies investigate how different extrinsic and intrinsic motivational factors can affect the participation of students in the DescribeIT socially responsible crowdsourcing project. In order to create a sustainable crowdsourcing project to describe images in digital learning resources for blind and partially sighted students, we need to look beyond the motivation of monetary rewards and explore what other motivations, either extrinsic or intrinsic, might work. In addition, in each study a comparison will be made between students' self-reported motivations and their actual behaviour to investigate the magnitude of the social desirability bias in reporting motivations to participate in socially responsible crowdsourcing projects.

The first study in the chapter *(Study 2)* investigated the effect of the intrinsic motivator of students' sense of altruism, and the extrinsic motivator of improving students' study skills. The second study in the chapter (*Study 3)* investigated the effect of the intrinsic motivator the benefit that blind and partially sighted students have from the project and the extrinsic motivator of increasing students' understanding of the teaching materials. Finally, the third study (*Study 4)* recruited students from the Department of History at the University of York, a department known for its excellent record of accepting blind and partially sighted students. This study investigated whether students in a department known for its involvement with visually disabled students, who might be aware of its reputation, would be more motivated to participate in the DescribeIT Project than students in other departments.

In Studies 2, 3, and 4 students' motivation when they are engaged with the image description task was measured using the Situational Motivation Scale (SIMS) (Guay et al. 2000) to assess their motivation while actually doing the task. In Studies 2, 3 and 4, as explained in Chapter 2 (See section 2.8), it was hypothesised that people who are more altruistic, that is more considerate of others, are more likely willing to participate in a socially responsible crowdsourcing project to support blind and partially sighted students. Therefore, students' general altruistic behaviour was measured using the Self-Report Altruism Scale (SRA) (Rushton et al. 1981). In addition, as people's attitude towards disabled people could potentially affect their behaviour in a crowdsourcing project to support disabled students, students' attitude towards disabled people was measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994).

5.2 Preliminary studies

Before conducting Studies 2, 3 and 4, it was necessary to conduct two preliminary studies. The first preliminary study established a classification of the images found in the digital learning resources to be used in the studies, to guide the sampling of images for Study 2 and further studies. The second preliminary study validated the intrinsic and extrinsic motivation instructions to be used in the studies, to ensure that they were understood as intended.

5.2.1 Preliminary Study A: Classification of the images used in Department of Computer Science digital learning materials

5.2.1.1 Introduction

The aim of this study was to develop a classification of the images used in digital learning resources that were to be used as stimuli in the programme of research studies. The classification would be used to ensure that the selection of images used in the main studies cover a representative range of images used by lecturers in their digital learning resources.

The particular digital resources used in this study were PowerPoint slide packs used by a range of lecturers in the Department of Computer Science.

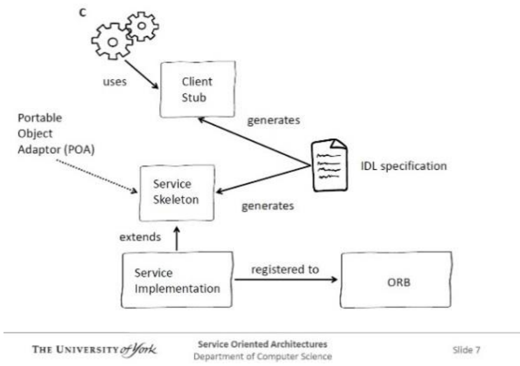
5.2.1.2 Method

The study was conducted with students on the Masters programme in Human-Centred Interactive Technologies (HCIT) as part of one of their practical. A total of 13 students participated, 4 men and 9 women. The study used a card sort method, asking students to group the images into sets they found similar. More specifically, students were asked to do an open card sort task. Students were given a pack of 30 slides each with an image from lectures on a range of MSc modules in the Department of Computer Science at University of York. Students were asked to work individually to place the images into groups, as many or as few groups as they wished, depending on what they thought the type of image was. They were asked to provide a text label describing each group.

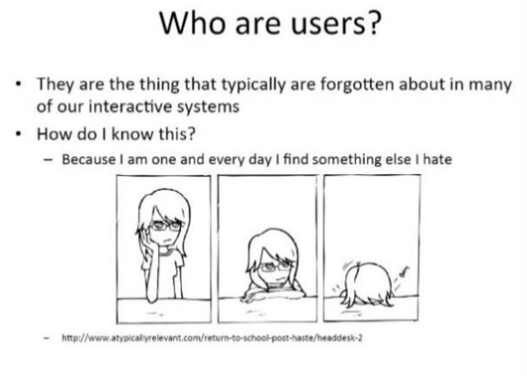
5.2.1.3 Results and Discussion

A cluster analysis was undertaken this showed six image categories, including: Graphs and charts, Abstract diagrams, Box and line diagrams, cartoon, Pictures of technology, and screenshots. Table 5.1 shows examples of images within the categories.

Table 5.1 Example of images categories produced from the cluster analysis

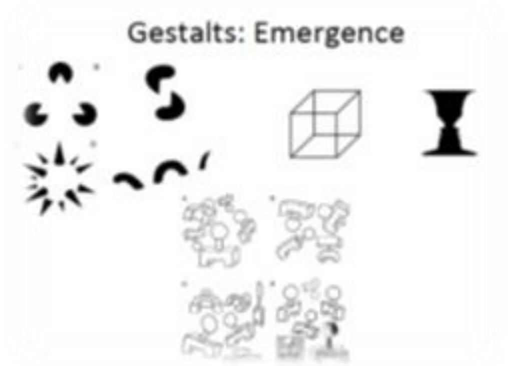


Box and line diagram

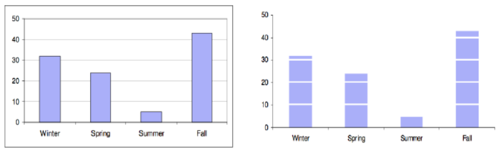


Cartoon

Screenshot



Abstract diagram



Graphs and charts.

Pictures of technology.

5.2.2 Preliminary Study B: Validation of the motivational instructions

5.2.2.1 Introduction

A second preliminary study was conducted to validate the intrinsic and extrinsic motivations in the recruitment emails and project information pages to be used in

Studies 2 and 5, it was felt it was not necessary to validate the intrinsic and extrinsic motivations of the other studies reported in this thesis.

5.2.2.2 Method

The instructions were generated by one group of participants and then validated by another groups of participants. In the first group, four PhD students, two from the Department of Education, and two from the Department of Electronics were asked to generate textual instructions to promote the chosen intrinsic and extrinsic motivational factor. Participants were given the definitions of intrinsic and extrinsic motivational factors from Dice and Rayn Self-Determination Theory (See Chapter 2, section 2.4.2), and were asked to word the instructions to emphasise the underlying reason of participation in the DescribeIT project in each motivational condition.

The generated instructions were then given to the second group of participants, eight PhD students, one research assistant and one lecturer in the Department of Computer Science. Participants were given a copy of the recruitment email and several options of instructions to promote participants' altruistic behaviour and several options to promote improving participants own skills. Participants were asked to rank the 4 extrinsic motivation instructions options from the most preferred to the least preferred option (i.e. 1 for their most preferred, and 4 for the least preferred), and to rank the 6 intrinsic motivation instructions options from the most preferred to the least preferred option (i.e. 1 for their most preferred, and 6 for the least preferred).

5.2.2.3 Results and Discussion

Participants of the first round generated a total of 10 instructions, six for altruism and four for improving one's skills. Consequently, participants of the second round ranked these instructions. Based on the highest ranked instructions, the textual instructions used through the recruitment emails and project information page were chosen.

5.3 STUDY 2: The Effect of Intrinsic vs. Extrinsic Motivational Factor 1

5.3.1 Introduction

Study 1, reported in Chapter 4, investigated what students thought would motivate them to participate to the DescribeIT project. The present study investigated UK students' actual participation in the DescribeIT project to provide image descriptions for digital learning resources. More specifically, based on the results of Study 1 and Phase 1 of the present study, the study investigated the effect of the intrinsic motivator of students' sense of altruism, and the extrinsic motivator of improving students' study skills. The aim was to investigate if these motivational factors would motivate students to participate in the DescribeIT project. In addition, students' self-reported motivations will be compared to their actual behaviour to measure the social desirability bias in the self-report data.

The Situational Motivation Scale (SIMS) (Guay et al. 2000) was employed to assess students' motivation while actually doing the task. In addition, students' general altruistic behaviour and attitude toward disabled people were measured by the Self-Report Altruism Scale (SRA) (Rushton et al. 1981) and the Interaction with Disabled Persons Scale (IDP) (Gething, 1994) respectively.

5.3.2 Method

5.3.2.1 Design

This study investigated how different motivational factors might affect the motivation of university students in a socially responsible crowdsourcing project. The study consisted of two phases, as follows:

Phase 1: investigated what UK university students think would motivate them to participate in a socially responsible crowdsourcing project, DescribeIT, to describe images for their blind and partially sighted peers. This phase used an online questionnaire (the same online questionnaire used in Study *1, see* 4.2.3).

Phase 2: investigated the behaviour of the students when they were invited to participate in the DescribeIT project to describe images for their blind and partially sighted peers. There were two conditions: intrinsic and extrinsic motivation.

Phase 2 of the study used a between-participants design. The independent variable was the motivation given to the students to participate in the project. Motivations were provided in the instructions in the recruitment email and on the project information webpage. There were two conditions: an intrinsic motivation condition (InMot) and an extrinsic motivation condition (ExMot). The particular intrinsic and extrinsic motivators were chosen based on the results of Study 1 and the results of Phase 1 of the present study. The intrinsic motivator, sense of altruism, was rated the highest of the 12 motivational factors and was the most frequently mentioned intrinsic motivational factor in students' open-ended questions both in Study 1 and the present Study. The extrinsic motivator, improving study skills., was rated the third highest extrinsic motivator. However, the two highest extrinsic motivators, monetary reward and academic credit were omitted because (1) the point of the studies was to investigate non-financial motivational factors, so monetary reward was not appropriate; (2) I would not been able to reward students with academic credit.

In Phase 2, students were randomly assigned to one of the two motivation conditions. The main task in Phase 2 was to create descriptions of images in digital learning resources. Students were given a set of guidelines about how to create descriptions of images suitable for blind and partially sighted students.

The dependent variables were the number of students who participated in the study from the total population of students asked to participate and the number of images described by each student who participated. Further dependent variables were:

Students' overall level of altruism measured using the Self-Report Altruism Scale (SRA), Students' overall attitude toward disability measured using the Interaction with Disabled Persons Scale (IDP) and students' motivation while doing the tasks measured using the Situational Motivation Scale (SIMS). See Chapter 2, section 2.8 for full description of the rationale behind measuring these variables.

For the purposes of the current programme of research, some of the items of the SRA were thought to be inappropriate for the target populations of students in the UK. For example, "I have helped push a stranger's car out of the snow". As the SRA consists of a single factor, four items from the scale were dropped, and the remaining 16 were used. When comparisons were made with previous research, scores were scaled appropriately.

5.3.2.2 Participants

For Phase 1, a total of 73 M.Sc. students in the Department of Computer Science at the University of York were invited to take part in the study. 22 students responded (a response rate of 30%), including 7 women (32%) and 15 men (68%). Their age range was from 21 to 38 years old, with a mean age of 25.8 years (SD=4.6). The University of York is a very international institution, and while some students were from the United Kingdom (23%, 5), a number were from other countries. 23% (5) were from China, the remaining students (54%) were spread relatively evenly across 9 other countries: Bangladesh, Bulgaria, Chile, Cyprus, Greece, Kenya, Malta, Mexico, and the Netherlands.

All students reported using social media, such as Facebook, Twitter, and LinkedIn except for one student who reported not using social media. However, only 3 students (14%) reported participating in crowdsourcing projects, and only a few times a year.

Crowdsourcing project they had been involved in included developing patches for Android and Fold It[4].

As an incentive to complete the questionnaire, students were entered into a prize draw for one of 10 Amazon vouchers worth £10 each.

For Phase 2, the same population of 73 MSc students was invited again to take part in the study. 8 students participated (response rate of 11%), including 5 students in the intrinsic motivation condition and 3 students in the extrinsic motivation condition. However, none of the students completed the IDP, SRA, and SIMS questionnaire or the demographic questionnaire, so no demographic data was about these eight students.

5.3.2.3 Materials

Phase 1: the questionnaire was the same one used in Study 1 (see Section 4.2.3 for details).

Phase 2: The motivational instruction for the InMot condition was: *The care you provide to others can only give you good things in return. Please help us improve the accessibility of images presented in PowerPoint slides*. The motivational instruction for the ExMot condition was: *Participating in this project will improve your skills to make websites, apps and software accessible to people in the future. Following the image description guidelines provided with this system will help you learn how to provide good image descriptions. In addition, this can help you improve your skills in how to communicate clearly in computer science using images*.

These instructions were included in both the recruitment email and the DescribeIT project information page.

The online questionnaire of Phase 2 (See appendix C.1.3), included four sections:

The *Situational Motivation Scale (SIMS)* (Guay et al. 2000) which includes four subscales: Intrinsic Motivation, Identified Regulation, External Regulation, Amotivation. The *Self-Report Altruism Scale (SRA)* (Rushton et al. 1981) and the *Interaction with Disabled Persons Scale (IDP)* (Gething, 1994), Maclean and Gannon's (1995) sub-scales were used which measure "Discomfort" and "Sympathy" toward people with disabilities. See Chapter 2, section 2.8 for full description about these

[4] https://fold.it/portal/

scales. *Demographic questionnaire:* collected information such as students' age, gender, and previous experience with crowdsourcing.

5.3.2.4 The DescribeIT Project

The DescribeIT project used in the present study was presented in Chapter 3, section 3.2. As there were two conditions with different instructions, two projects were created. The projects were identical in every aspect except for the instructions given to promote the InMot and ExMot factors (See 5.3.2.3 for these instructions). To help students create good image descriptions, the project information page provided guidelines on how to describe images for blind and partially sighted people and an example description of a typical image [See appendix A.1]. Students were told that they could describe as much or as few as they wish, and that they could log into the project on different occasions to describe images.

The project consisted of 60 PowerPoint slides, from different Computer Science courses, each containing an image. Figure 5.1 shows some examples of the images used in the project. The range of images used covered all types of images used by lecturers of Computer Science PowerPoint presentations (details on section 5.2.1). Each slide was available to be described by four different students (i.e. 240 descriptions could be produced).

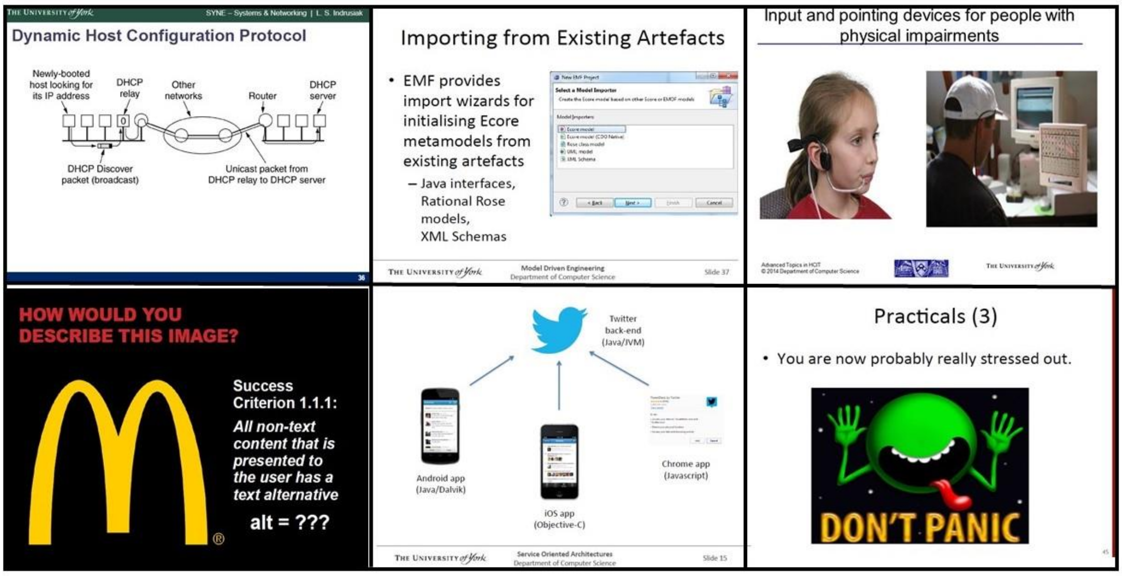


Figure 5.1 Some examples of the images used in the DescribeIT project

5.3.2.5 Pilot study

A pilot study was conducted to check the manipulation of motivations and to check for any technical problems with the crowdsourcing project. Participants were five PhD students of the University of York, three from the Department of Computer Science, one from the Department of Education and one from the Department of Electronics. Participants were asked to take part in the Phase 1 online questionnaire. A few days later they were asked to participate in the DescribeIT project. The results of the pilot suggested some minor changes in the layout of Phase 1 questionnaire. In addition, in the crowdsourcing project the progress bar which shows how many images a student has described and how many still available for descriptions was removed, as four participants thought this could hinder them from participation if there were a lot of images to describe. Five images which were reported to be difficult to describe by more than one participant were replaced with other images from the same category. Lastly, some cosmetic changes were made to the project information page. Participants in the pilot study did not participate in the main study.

5.3.2.6 Procedure

For Phase 1, a recruitment email was sent out to all currently enrolled M.Sc. students in the Department of Computer Science (see Appendix C.1.1). A reminder email was sent five days later. The online questionnaire was available for two weeks.

For Phase 2, two different emails (see Appendix C.1.2) were sent to the same population of M.Sc. students targeted in Phase 1 of the study to take part in Phase 2 of the study, one to promote the intrinsic motivational factor and another to promote the extrinsic motivational factor. Two reminder emails sent five days later each. The project was available for four weeks.

5.3.2.7 Data Analysis

Some of the data collected were not normally distributed, except for four variables which were normally distributed, namely: being connected, fun/entertainment, contributing and social recognition. Thus, it was decided to use nonparametric tests for all analyses. Parametric tests for the normally distributed data showed the same results found with non-parametric tests.

Qualitative analysis of the two open-ended questions about students' motivations to take part in the DescribeIT and the open-ended question about students' concluding thoughts were carried out using content analysis (See 4.2.5 for more details).

5.3.3 Results

Phase 1: Self-Reports of Motivation

Students were asked to rate their likelihood and other students' likelihood to participate in the DescribeIT project, on 7 point Likert items (1=Not at all likely, 7=Very likely). Table 5.2 summarizes the means, medians and standard deviations for these two questions, and Figure 5.2 shows the mean ratings for both questions. Students rated their likelihood of participation with a mean rating of 4.6 (*SD=1.8*), and rated the likelihood of participation of their peers with mean rating of 3.8 (*SD=1.4*). One sample Wilcoxon signed rank tests were carried out to see if students' likelihood of participation rating for themselves and other students are statistically different from the neutral midpoint rating of 4 on the 7-point rating scale. The results showed that both ratings were not statistically significantly different from the midpoint of the scale. A Wilcoxon signed-rank test showed that students rated the likelihood of themselves participating in the project statistically significant higher than the likelihood of their peers participating in the project (Z = -3.02, $p < 0.005$).

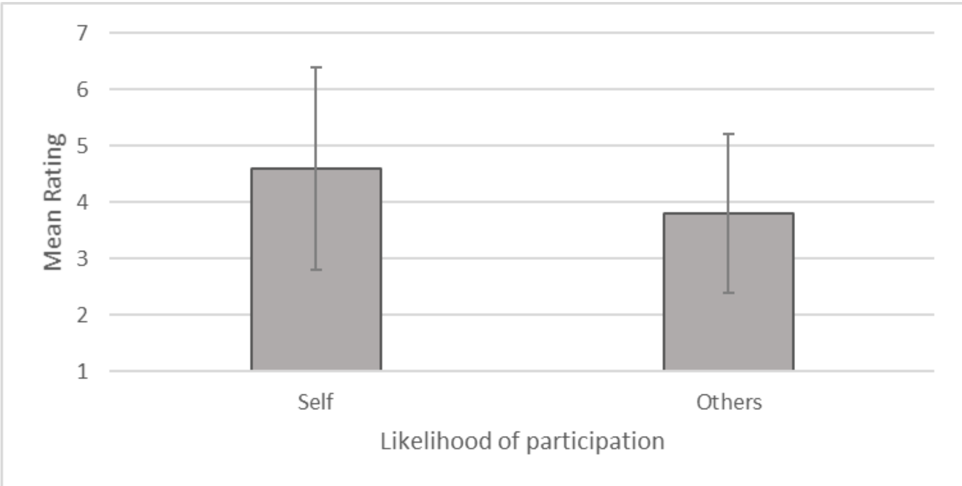


Figure 5.2 Mean ratings and standard deviation on the likelihood of own participation and of other students in the crowdsourcing project

Table 5.2 Mean and median ratings, standard deviation and results of one-sample Wilcoxon test on the likelihood of own participation and other students in the crowdsourcing project

Likelihood of participation	**Mean**	**Median**	**SD**	**W**	***P***
Self	4.6	5	1.8	1.4	0.160
Others	3.8	4	1.4	- 0.8	0.431

In the open-ended question students noted a number of factors which would motivate them and other students to participate in the DescribeIT project. This includes the motivational factors presented in Table 5.3. The most mentioned factors were financial rewards and wanting to help others.

Seven students noted the importance of financial rewards for them to participate to the DescribeIT project, in comparison to nine students who thought that this factor would motivate their peers to participate. Some of the students who thought that money would motivate them said:

Some type of reward (P1)

Monetary compensation, shopping vouchers/rewards (P16)

Whereas, some of those who thought it would motivate their peers said:

Money. Fame. A little bit of wanting to help others (P6)

free food e.g. pizza (P9)

Perhaps financial reward (P15)

Even when fun was mentioned it was associated with other incentives, for example, a participant said:

Something fun and with an incentive. (P7)

Six students thought that wanting to help others would motivate them. However, only four students thought it would motivate their peers. A student said:

I like to help other people (P3)

Students who thought that wanting to others would motivate their peers said:

Aiding other students (P17)

The basic motivation is that this activity is for a very good cause and aiming to help people (P5)

Improving students' understanding of the teaching materials was thought by five students to be a motivator to them, and three thought that would be a motivator for their peers. Some of the students' thoughts were:

Trying to understand and explain the images. To learn what the images would be teaching, i.e. free lectures. (P6)

If it helped some way towards my studies, for example a better understanding of a subject area (P15)

When one of the students thought of his peers, he said:

Something that would be advantageous to their study (P12)

Seeing how the project could help blind and partially sighted students was thought by two students to be a motivator for them and their peers. For example, a participant said what would motivate him and peers is:

Being able to see from the people it really benefits from the application (P4)

Table 5.3 Categorization of factors which students thought would motivate themselves and other students to participate in the crowdsourcing project (from open-ended question)

Motivational factor	**Self** (N, %)	**Other students** (N, %)
Financial rewards (e.g. cash, vouchers)	7 (32%)	9 (41%)
Wanting to help others	6 (27%)	4 (18%)
Knowing that participating could help understanding the teaching materials	5 (23%)	3 (14%)
Seeing how this could help blind students	2 (9%)	2 (9%)
Knowing a blind person	2 (9%)	1 (5%)
Lecturers asking them to participate	2 (9%)	2 (9%)
Fun to do	2 (9%)	1 (5%)
University recognition	1 (5%)	1 (5%)
Easy to do	1 (5%)	1 (5%)
Being interested in the accessibility area	1 (5%)	0 (0%)
To show off	0 (0%)	1 (5%)
Looking for fame	0 (0%)	1 (5%)
Advertising the project	0 (0%)	1 (5%)

Similarly, two students thought that if they were asked by their lecturers to participate to the project, that would motivate them and their peers. One of the students' thoughts on this factor, were:

If the lecturer asked me to and explained why it is needed (P12) [Self]

Not having a choice. (as in, the lecturer told them they have to) (P12) [Others]

When considering his comments, we can notice that the student used the words *asked* and *explained* by the lecturer why the project is needed for him to participate. However, when he thought of his peers he said *not having a choice* and *they have to*.

Knowing a blind student was thought by two students to be a key motivator to them, whereas only one students thought that would motivate his peers. The following is how a student thought this factor motivate himself and his peers:

If I had relatives or friends that are visually impaired, then it would surely motivate me (P21) [self]

Same as my previous answer. (P21) [others]

The fun of doing the description task was thought by two students to be a motivator, in comparison to only one student who thought it would motivate other students:

Sounds fun (P8)

Whereas the only student who thought fun could motivate peers said:

Something fun and with an incentive. (P7)

One student thought that the image description task would be a difficult one to do, she also thought that this could negatively affect her peers' participation as well:

Sounds hard to me…. If it's hard, people are not much motivated (P11)

One student mentioned the social recognition students might receive from the university for participation, which would motivate him and his peers.

A formal recognition from the University that I helped in that project would be nice. (P19)

Being interested in the accessibility area was only thought by one student to be a motivator, he said:

I am interested in accessibility technologies (P3)

Table 5.4 Mean and median ratings, and standard deviation of 12 motivational factors with results of one-sample Wilcoxon Signed Rank tests against neutral midpoint

Motivational Factor	**Mean Rating**	**Median Rating**	**SD**	**W** ***p***
Your sense of altruism, wanting to help other students (Altruism) (IN[1])	5.2	5.0	1.5	3.01 0.003
Being paid for your efforts (Money) (EX[1])	4.8	4.0	1.9	1.80 n.s
Knowing that you are contributing to a large project (Contributing)(IN)	4.6	5.0	1.6	1.43 n.s
Getting academic credits (Academic credits) (EX)	4.3	4.0	2.3	0.48 n.s
Improving your academic skills (Improving skills) (EX)	4.3	4.0	1.6	0.67 n.s
The fun and entertainment of the activity (Fun/Entertainment) (IN)	4.0	4.0	1.8	0.09 n.s
Being connected with other students on your course (Being connected) (EX/IN)	3.9	4.0	1.8	-0.39 n.s
Enhancing your job opportunities in the future (Job opportunities) (EX)	3.8	4.0	2.2	-0.78 n.s
Being in a competition with other students (Competition) (EX)	3.1	3.0	1.8	-2.19 0.037
Drawing attention to your skills (Attention) (EX)	3.1	3.5	1.7	-2.11 0.036
The social recognition you would receive (Social recognition) (EX)	3.0	3.0	1.5	-2.54 0.011
To pass the time (To pass time) (EX)	2.6	2.0	1.7	-3.02 0.003

1. IN=Intrinsic motivation, EX=Extrinsic motivation

Interestingly, when students reported what they thought would motivate other students to participate in the DescribeIT project, they put more emphasis on rewards (different form of financial rewards) than when reporting about themselves. In addition, new motivational themes were reported that would motivate only peers, such as advertising the project, looking for fame and showing off. For example, a student said:

if output is good to show off (P10)

On the second part of the questionnaire students were asked to rate whether each of 12 factors would motivate them on 7 point Likert items (1= Not at all to 7=Very mu ch). The mean ratings are shown in Table 5.4 and Figure 5.3. Students rated altruism to be the highest motivational factor to participate in the image description activity. Whereas, to pass time was the least rated motivational factor.

One-sample Wilcoxon Signed Rank tests were conducted to test whether students' ratings of the 12 motivational factors were statistically different from the neutral midpoint rating of 4 (See Table 5.4). Only five motivational factors were significantly different from the neutral midpoint. The results showed that students were positively motivated by the intrinsic factor of sense of altruism (*p< .005)*. However, the median rating was significantly below the neutral midpoint for four factors: pass time (*p < .001),* social recognition*, (p< .05),* attention (*p< .05),* and competition (*p< .05).*

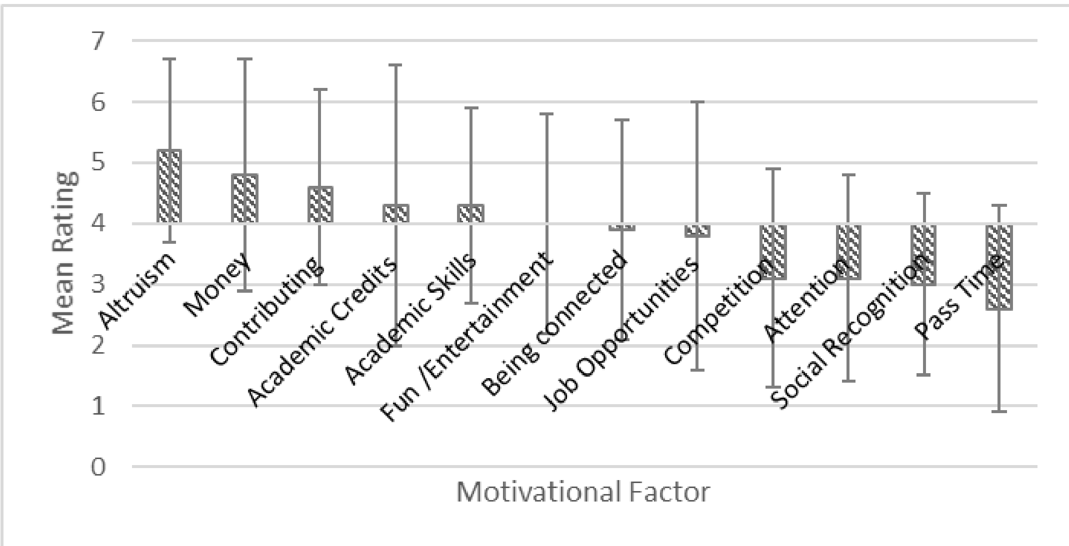


Figure 5.3 Mean ratings and standard deviation for the 12 factors as motivations for self against the neutral point on the rating scale

The following sections present more detailed results for each factor (more detailed examples of students' ratings and comments about each factor can be found in Appendix C.1.4). The motivational factors presented from the most important to the least important, as documented in Table 5.4

Students were generally motivated by their sense of altruism. However, students' comments on the motivational factor (See Table C.1, in Appendix C.1.4) showed that students who were less motivated by this factor were concerned about having spare time to contribute to the project. Students though that participating in the project would make them happy.

The second highest motivational factor was monetary reward, yet this factor was not significantly different from the midpoint of the scale. Students who were positive about this factor thought that extra cash is always good, even if they were altruistically motivated to participate to the project (See Table C.2). Similar to students' comments on the altruism factor, students seemed to have concerns about having the time to participate.

Students were also on average neutral on their feeling about knowing that they are contributing to a large project. Students comments (See Table C.3) showed that they did not think of this motivational factor as a key factor to motivate them to participate. However, those who thought that this factor would motivate them thought that being involved in a large project meant that their contribution would impact large group of people, which would make them feel that their contribution was more valuable.

Students were also on average neutral in their feeling about getting academic credits mainly because they did not feel it is right to get academic credits out of their participation in the DescribeIT project (See Table C.4). However, the idea of getting academic credits which potentially could enhance students' career sounded appealing to some students.

Similarly, students were on average neutral in their feeling about improving their academic skills. As students' comments explained (See Table C.5), this was due to the fact that students could not see how describing images to blind and partially sighted students would improve their skills. However, those who could see how this task would potentially improve their skills were motivated by this factor.

The description task was not thought by students to be a fun task to do. This was hinted at in some of students' comments that the description task could be a difficult one to do, especially if the images are complex. It was also reflected by motivated students' comments (See Table C.6), as for those students the idea of doing the description task would probably be challenging, which seems to be what would make any task fun to do.

Being connected with other students did not seem to be an important factor to students. Students thought that there are many other ways to be connected with their peers (See Table C.7). Interestingly, those who were positive about this motivational factor thought that participating in the project might connect them with disabled students.

Students were on average neutral in their feeling toward the possibility of the task enhancing their job opportunities in the future. As the comments showed (See Table C.8), this was mainly because students were not interested in pursuing a career in a field relevant to accessibility.

Students were on average not motivated by the idea of being in competition with their peers to describe images for blind and partially sighted students. For most students competing with their peers in helping disabled students was not appealing to them (See Table C.9). However, even those who were in favour of this motivation had concerns regarding the quality of the work that could be produced as result of being in competition.

In addition, students were not motivated by the idea of drawing attention to their skills. Perhaps this related to the fact that students were not particularly interested in the accessibility area. See Table C.10 for some of students' comments on this factor.

Students were also not motivated by the social recognition factor. As shown in students' comments (see Table C.11), students thought their participation should come from wanting to help others rather proving themselves to others. However, the effect of social norms was hinted at in some of students' comments on this factor. For example, participating in a project to support disabled peers is socially appealing, in contrast not participating in a project to support your disabled peers may seem not socially acceptable.

Lastly, students were not at all motivated by the idea of passing the time through describing images. As Table C.12 shows, students thought there are many other ways to pass their time in more enjoyable way than describing images.

At the end of the questionnaire students were asked for any final thoughts of what would motivate students to participate in the DescribeIT project. Students gave some interesting ideas, including:

- Include a ranking system or badges: *Maybe a ranking with high scores could help … virtual prizes like badges.* (P3)
- Including other gamification aspects: *Organise the activity by levels of difficulty, so users may become players advancing through several stages.* (P3)*, Integrating animation would help.* (P7), *Sound?* (P8)
- Providing feedback to participants: *Provide update on the success rate* (P11)
- Targeting undergraduate students: *You'll need to find the few who are just starting out at uni[versity] and need to build their social portfolio. I'm perhaps the wrong audience because I'm a Masters student with a pre-defined set of interests now* (P16)
- Introduce potential participants to blind students: *allow them to make relationships with the people they are helping and it will become all that much more important in their agenda* (P2)

Phase 2: Behaviour in the DescribeIT Crowdsourcing Project

Overall, 8 students participated in Phase 2 of the study, 5 (62.5% of the eight participants) students in the InMot condition and 3 (37.5%) students in the ExMot condition.

Out of the eight students only five had participated in Phase 1 of the study, thus it was not possible to conduct a statistical analysis comparing the results of Phase 1 with those of Phase 2 as had been planned.

In the InMot condition students viewed 207 images but described only 83 images, meaning they viewed but failed to describe 124 images. The number of images described ranged from 2 to 25 images per student, a mean of 7.8 (SD= 10.0) descriptions per student. On the other hand, the ExMot students viewed only 21 images, and described only three images, meaning they viewed but failed to describe 18 images. Each student described only one image.

Unfortunately, none of the 8 students responded to Phase 2 questionnaire which included the SIMS, SRA and IDP measures.

5.3.4 Discussion and Conclusions

The study aimed to investigate the intrinsic motivational factor, altruism and wanting to help others, and the extrinsic motivational factor, improving one's skills. The study also aimed to investigate what MSc. students think would motivate them to participate in the DescribeIT project in comparison to their actual behaviour in participating in the project. Building on Study 1, presented in Chapter 4, the results of this study provide further insight into what students say would motivate them to participate in a socially responsible crowdsourcing project and a first insight in their actual behaviour when participating in such a project. For results summary see Tables 5.12 and 5.13 end of this chapter.

The results of Phase 1 of the present study showed that students were not significantly positively motivated to participate in the project. Students also thought that their peers would be neutral in their motivation to participate in the project. This first finding contrasted with the findings of Study 1, in which students thought that they would be positively motivated to participate in the DescribeIT project. Although in the present study the overall median was above the midpoint of the rating scale, students were split equally below and above the midpoint of the scale resulting in no significant deviation. However, students' rating of the likelihood of other students participating in the project agreed with the findings of Study 1. In line with the previous study, in this study students rated their likelihood of participating significantly higher than the likelihood of other students participating. This could be due to what social psychologists call the "fundamental attribution error" (Ross and Nisbett, 2011) which proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits. Hence, when students were asked to predict their own participation, they thought of the specific situation of helping disabled students, but when asked to predict the behaviour of other students, they thought of the general helpfulness of other students, not in the context of helping disabled students. A similar pattern was noted when students commented on what would motivate them and what would motivate other students to participate in the project.

In terms of what would motivate themselves and others, students said they would be particularly motivated to participate in the crowdsourcing project by financial rewards and their sense of altruism. This result is in line with the findings of Study 1, in which students thought they would be particularly motivated by these two factors. Students of the present study also reported that knowing that their participation could help them in understanding the teaching materials would motivate them and their peers to participate in the DescribeIT project. Although this factor was also reported in Study 1, it was more important in the present study in comparison to Study 1. In addition, students noted other motivational factors that could motivate them to participate, such as seeing how the project is useful to blind and partially sighted students, knowing a blind person, if their lecturers asked them to participate, and the task being fun to do, among others. However, all these factors were reported by less than 10% of the students.

Moving to students' ratings of the 12 motivational factors, students reported that they were significantly motivated by their sense of altruism and wanting to help others. This finding is in agreement with Study 1 and previous research (e.g. Kuznetsov 2006, Oreg and Nov 2008). However, the finding of the ratings in the present study did not completely match students' comments on what would motivate them to participate to the project. For example, students rated altruism the highest motivational factor, whereas in the open-ended question it was the second highest mentioned motivational factor.

More interestingly were the results of monetary reward factor, where students' rating did not show a significant difference from the neutral midpoint of the scale. This finding disagrees with the findings of the previous study, and also with previous research by Lakhani et al. (2007) and Brabham (2008, 2010), in which the importance of monetary reward as a self-reported motivation was reported. Whilst it can be understandable that the finding of this study is different from those of Lakhani et al. (2007) and Brabham (2008, 2010), as the context of the DescribeIT project examined in this study is not similar in nature to the crowdsourcing context examined by Lakhani et al. (2007) and Brabham (2008, 2010). Nonetheless, the fact that the finding of Study 1 is not in agreement with the finding of the current study is interesting, as both studies explored students' motivation to participate to the same crowdsourcing project to support blind and partially sighted students. Furthermore, in the present study students' rating to the monetary reward factor did not match students open-ended questions' answers, which showed that this factor was the most mentioned factor by students. However, it is worth noting that the number of students participating in this study is relatively small in comparison to the previous study (22 in Phase 1 of the current study compared to 271 in the previous study). Also, the previous study targeted undergraduate students in comparison to postgraduate students in this study, who might be comparatively more materially secure. Therefore, this particular finding requires further investigation.

The results of Study 1 found that students were significantly motivated by the idea of getting academic credits for their participation in the DescribeIT project. However, the present study found a neutral self-report rating for getting academic credits, which students felt was not appropriate. When students were asked in the open-ended question about motivations none mentioned getting academic credits.

Consistent results were found between the present study and Study 1 with regards to the motivational factor of improving one's skills, as in both studies students thought they were neutral about this factor. This finding contrasts with Brabham's (2008, 2010, 2012) findings, which as mentioned in the Discussion of Study 1 could be due to the fact that improving artistic skills may not be the same as improving academic skills. However, although students did not mention explicitly the motivational factor of improving one's skills in their answers to the open-ended question of what would motivate them and others, they thought that the idea of increasing their understanding in the teaching materials would motivate them and other students to participate.

Similar to the results of Study 1, the results of this study suggest that students thought they were neutral in their motivation by the fun and entertainment of the activity, which is contrary to some previous studies (Brabham 2008, 2012, Hossain 2012, Kaufmann, et al. 2011). This is another indication that students did not perceive the description task as fun task to do, perhaps because students thought of it as difficult task to do.

Previous research (Kuznetsov 2006, Brabham 2010) emphasized the importance of the community and being connected to other participants in crowdsourcing projects. However, the results of the present study along with the results of Study 1, both suggest that being connected to other students was not thought by students to be a important motivation to participate in the project. Interestingly, however, some students thought the project could be a way of connecting with disabled students.

The findings of Study 1 were in line with the findings of the present study regarding competing in relation to helping blind and partially sighted students, in that this motivational factor was not appealing to students in both studies. In addition, students in the present study and in Study 1 were not motivated by drawing attention to their skills. This was mainly because students were not particularly interested in the accessibility area, thus participating in this project would not draw attention to their actual skills in this area. Consistent results were also found about the motivation of passing the time in both the previous study and the current study, students did not think that describing images for blind and partially sighted students would be a task they would do to pass the time. This was also reported in previous research by Brabham (2008) when studying iStockphoto members.

The unexpected results of Phase 2 of the study did not allow for any statistical analysis, however, the fact that the number of students of the intrinsic motivational factor condition were almost twice as the number of the extrinsic motivational factor condition is an interesting finding. Nonetheless, the fact that students' ratings of the likelihood of participation in Phase 1 were not significantly different from the neutral midpoint of the scale means that this particular group of students were not generally motivated to participate to the DescribeIT project.

To conclude, the results of this study present further insight into students' self-report of motivations to participate in the DescribeIT project and first insight into students' actual behaviour when asked to participate in the DescribeIT project. Students self-report of the likelihood of participation showed that students were not generally positively motivated to participate to the DescribeIT. This was reflected in students very low participation rate in Phase 2 of the study. When students were asked to participate in the online questionnaire (Phase 1) with the chance of a reward, 22 students participated. Whereas, when students were asked to participate in the crowdsourcing project (Phase 2) with either an intrinsic motivation or extrinsic motivation (which was not a monetary reward) only 8 participated. In addition, students' ratings of the 12 motivational factors did not completely match their answers to the open-ended questions, especially to the two leading motivational factors: altruism and money.

However, these results need to be considered with caution for several reasons. Firstly, the sample size of this study is small. It was hoped that a larger number of the MSc students would participate in the DescribeIT project. Secondly, the sample is consisted only of students studying Computer Science, who may not be typical of all university students in their attitudes and motivations towards crowdsourcing. Hence, the population of students will be expanded in the subsequent studies to include both undergraduate and postgraduate students from other Departments across the university, including the Theatre, Film and Television, and History Departments. Lastly, Phase 2 was conducted approximately two weeks before MSc students were due to submit their research projects. Probably because of this circumstance, this may account for the low participation in Phase 2. For this reason, I did not consider this to be a true representation of the participation rate, and further data collection will be carried out with students in the next cohort, in order to achieve a larger sample.

5.4 STUDY 3: The Effect of Intrinsic vs. Extrinsic Motivational Factor 2

5.4.1 Introduction

In line with Study 2 this study aimed to investigate the effect of an intrinsic and extrinsic motivational factors in students' participation to the DescribeIT project. More specifically, Study 3 aimed to investigate two motivational factors which emerged from the previous studies (1 and 2): knowing how blind and partially sighted students would benefit from the project and increasing students' understanding of their teaching materials. In addition, the relation between students' participation and their situational motivation, their attitude toward people with disabilities, and lastly to their self-report sense of altruism was also investigated. The self-report data were also compared to the actual behaviour to measure the social desirability bias in the self-report data.

Study 3 extended the population of students to include undergraduate students from both the Departments of Computer Science and Theatre, Film and Television. Not only did this mean that a larger pool of students was available, but also the population of the previous study extended to include students from the Department of Theatre, Film and Television, as students from different departments may differ in their attitudes and motivations towards crowdsourcing.

One reason for the low participation rate in Study 2 may have been the poor timing in relation to the students' schedule, being only several weeks before their most important deadline. Therefore, the timing of Study 3 was carefully chosen so students would not be too busy with submission deadlines or exams, to make sure that the timing of conducting the study did not suppress the participation rate.

To allow students to better understand the images within the context of the teaching materials, in this study the whole pack of PowerPoint slides was used in the DescribeIT project, instead of only presenting the slides contain images as in Study 2.

5.4.2 Method

5.4.2.1 Design

The study investigated an intrinsic factor and an extrinsic factor in students' motivations to participate in a socially responsible crowdsourcing project, the DescribeIT project to describe images for blind and partially sighted students. The intrinsic motivational factor (InMot) was knowing how blind and partially sighted students would benefit from the project and the extrinsic factor (ExMot) was increasing students' understanding of their teaching materials. These come up in Study 1, 2 and Phase 1 of the present Study. The intrinsic motivational factor was implemented by showing students a short video of a blind student talking about the problems he faces in accessing images in his digital learning materials. The extrinsic factor was implemented by a textual instruction both in the recruitment email and the project information page.

Similar to Study 2 this study consisted of two phases, as follows:

Phase 1: investigated how students perceive a range of motivational factors that would affect their participation and the participation of other students in a crowdsourcing project to describe images for their blind and partially sighted peers.

Phase 2: investigated the behaviour of the students in a crowdsourcing project to describe images for their blind and partially sighted peers. There were two conditions: intrinsic and extrinsic motivation. Conditions were randomly allocated to Departments, this does not necessarily mean that students were randomly assigned to groups. However, the allocation by departments was decided to avoid students from the same course talking to each other about the project, which could result in revealing the fact that there were different conditions.

The dependent variables were the number of students who will participated in the project and the number of images described.

Further dependent variables were:

Students' overall level of altruism measured using the Self-Report Altruism Scale (SRA), Students' overall attitude toward disability measured using the Interaction with Disabled Persons Scale (IDP) and Students' motivation while doing the tasks measured using the Situational Motivation Scale (SIMS)

(See Chapter 2, section 2.8 for full description of these three scales).

5.4.2.2 Participants

For Phase 1, 169 students were invited to participate in the study. The participants were undergraduate students recruited from two different departments (Computer Science and Theatre, Film and Television) at the University of York. A total of 26 students responded (15.4%), 5 women and 21 men. Their age range was from 18 to 25 years old, with a mean age of 19 years (SD = 1.8 years). One student reported that he is partially sighted. The majority of students were from the United Kingdom except for three students who were from Germany, Belgium and Australia.

All students reported using social media, such as Facebook and Twitter. However, only 9 (35%) reported participating in crowdsourcing projects before, including PlanetHunters[5], Eyewire[6], and Wikipedia[7].

As an incentive to complete the questionnaire, students were entered into a prize draw for one of five Amazon vouchers worth £5 each.

In the first round of Phase 2, no students participated. In a second round of Phase 2 (see section 5.4.5 for details), 4 students participated.

5.4.2.3 Materials

For Phase 1, the online questionnaire from Studies 1 and 2 was used (see section 4.2.3 for details).

For Phase 2, the projects presented digital learning resources containing images used by lecturers in the Departments of Computer Science and Theatre Film and Television at the time of conducting the study (i.e. the learning resources were from courses that students were taking).

The Phase 2 questionnaire consisted of four sections: Self-Report Altruism Scale (SRA) (Rushton et.al. 1981), Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000), Interaction with Disabled Persons Scale (IDP) (Gething, 1994), (See Chapter 1.6 for more details about each section) and demographic questions, as in Phase 1.

---

5 https://www.planethunters.org/#/

6 http://eyewire.org/explore

7 https://en.wikipedia.org/wiki/Main_Page

The video which was presented to students in the InMot condition was recorded with a blind second year undergraduate history student. The student talked about his day-to-day challenges as blind student, he also talked about how he usually overcomes some of these challenges. The video is available on the accompanying CD.

The recruitment email for the InMot condition included a link to 3-minute trailer of the video. On the project information page, a follow-up 10-minute video was available with a link to a longer 30-minute version of the video.

The motivational instruction for the ExMot condition was: *Describing the images in the Programming for Digital Media course will increase your understanding of the course materials. In addition, it will help you improve your skills in how to communicate clearly in Programming for Digital Media using images.*

5.4.2.4 The DescribeIT project

The DescribeIT project is presented in Chapter 3, section 3.2. Two versions of the project were developed, identical except for the motivational factor presented in the information page and the PowerPoint slides used. To help students create good image descriptions, the project information page provided guidelines on how to describe images for blind and partially sighted people and an example description of a typical image (see Appendix A.1). Students were asked to describe as many or as few as they wished, in as many sessions as they wished.

Different images were used in each condition. In the InMot condition, students were asked to describe images from the module Programming for Digital Media (Lecture 9: Image Processing), which they were taking at the time. In total, the project had 25 slides, 9 slides contained images. The aim was for each image to be described 30 times (approximate number of students in this module), which meant a total of 270 descriptions could be produced.

In the ExMot condition, students were asked to describe images from the module Human Aspects of Computer Science (Lecture 7: Prototypes), which they were taking at the time. In total, the project had 20 slides, 7 slides contained images. The aim was for each image to be described 100 times (approximate number of students in this module), which meant a total of 700 descriptions could be produced.

5.4.2.5 Pilot study

The two DescribeIT projects were piloted with two Electronic PhD students at the University of York. No technical problems were detected. Neither of the pilot participants took part in the main study.

5.4.2.6 Procedure

Recruitment emails were sent out to students inviting them to participate in Phase 1 of the study. Reminder emails sent out one week later. The online Phase 1 questionnaire was available for two weeks [See Appendix B.1.2].

For Phase 2 appropriate recruitment emails were sent out to students in the InMot and ExMot conditions to participate in Phase 2 of the study. The recruitment emails were the same except for the instructions section to promote either the InMot or ExMot condition [See Appendix C.2.1, C.2.2]. Both recruitment emails were sent out after students taken a class of web accessibility and assistive technology, which it was hoped would encourage students to participate in the DescribeIT project. Two reminders were sent out a week apart from each other.

5.4.2.7 Data analysis

The distributions of some of the dependent variables were not normal, thus nonparametric tests were used for all analyses.

Qualitative analysis of the two open-ended questions about students' motivations to take part in the DescribeIT and the open-ended question about students' concluding thoughts were carried out using content analysis (See 4.2.5 for more details).

5.4.3 Results

Phase 1: Self-Reports of Motivation

Since the students were from two different departments, Mann-Whitney tests were carried out to investigate if there was a difference between students' self-reports of motivations between the two departments, in terms of likelihood of participation and the 12 motivational factors ratings. The tests showed no significant difference in the results of Phase 1 between the two departments. Therefore, the results of Phase 1 for both departments will be reported together in the following sections.

Students rated their own likelihood to participate in the crowdsourcing project (on a 7 point Likert item, 1=Not at all, 7=Very likely), giving a mean rating of 5.0 (SD=1.2), whereas the likelihood that their peers would participate in the project they rated with a mean of 4.3 (SD=0.7). Figure 5.4 illustrates the mean ratings. One sample Wilcoxon signed rank tests were carried out to see if students' likelihood of participation rating for themselves and other students are statistically different from the neutral midpoint rating of 4 on the 7-point rating scale. Table 5.5 summarize the results. Students' rating for themselves ($W=3.2$, $p<0.001$) and for others ($W=2.1$, $p<0.05$) were both significantly above the midpoint of the scale. A Wilcoxon signed-rank test showed that students rated the likelihood of themselves participating in the project significantly higher than from the likelihood of their peers participating in the project ($Z = -2.6$, $p< 0.05$).

Table 5.5 Mean and median ratings and standard deviation on the likelihood of own and other students' participation in the DescribeIT crowdsourcing project

Likelihood of participation	**Mean**	**Median**	**SD**	**W**	***p***
Self	5.0	5.0	1.2	3.2	0.001
others	4.3	4.0	0.8	2.1	0.038

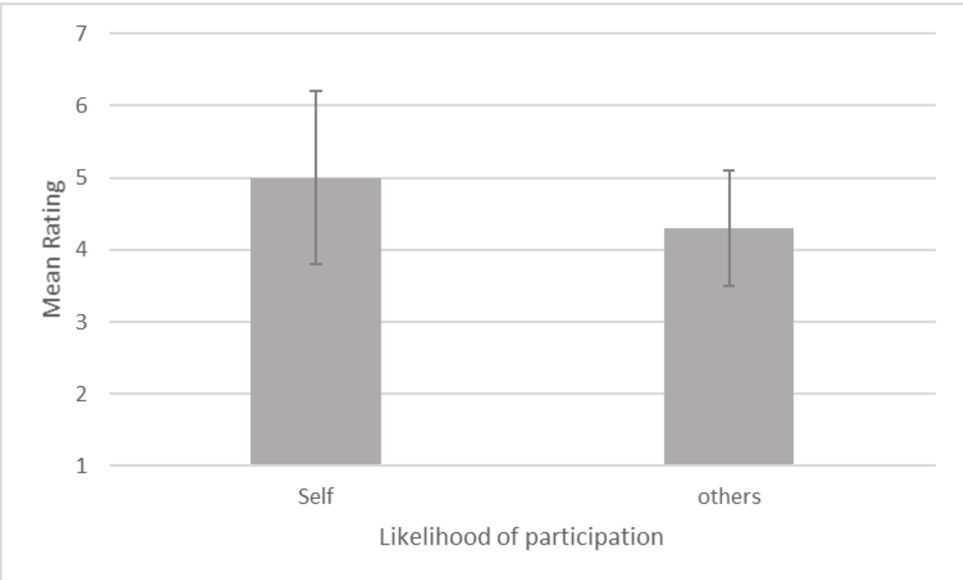


Figure 5.4 Mean rating and standard deviation of the likelihood of own and other students' participation in the DescribeIT crowdsourcing project

In the open-ended question about motivations for participating in the project, students reported a number of intrinsic and extrinsic factors that would motivate them and other students to participate. Table 5.6 shows the motivational factors reported, with the frequency of students' mentioning them. Overall, financial rewards were the leading self-report factor for both self and others, with most students thinking that this would motivate their peers to participate. In addition, students thought they would be motivated by their sense of altruism and wanting to help others, seeing how the outcome of their participation is useful to disabled students, and being in competition with other students. However, students thought that these motivational factors would be less important to their peers.

Ten students thought that some sort of financial rewards, such as Amazon vouchers or cash, would motivate them to participate to the DescribeIT project, in comparison to 19 students who thought that would motivate other students to participate. A student said:

As much as I hate to say it, external rewards would motivate me to put detail and consideration into the project more than if I was just volunteering. (P18)

When a student thought of other students, he said:

Financial incentives, e.g. vouchers or money. (P10)

Table 5.6 Categorization of factors which students thought would motivate themselves and other students to participate in the DescribeIT crowdsourcing project (from open-ended question)

Motivational factor	**Self** (N, %)	**Others** (N, %)
Financial rewards (e.g. vouchers, cash, prizes)	10 (38.5%)	19 (73%)
Wanting to help others	7 (27%)	2 (7.7%)
Seeing how their participation is helpful to disabled students	6 (23%)	1 (3.8%)
Being in competition with other students	2 (7.7%)	6 (23%)
The time required to participate not too long	2 (7.7%)	2 (7.7%)
Improving their CV	1 (3.8%)	1 (3.8%)
Receive a notification when his/her description being used	1 (3.8%)	0 (0%)
One's own disability	1 (3.8%)	0 (0%)

Seven students thought the intrinsic motivational factor of wanting to help others would motivate them. However, only two students thought that would motivate other students. When they thought of themselves and other students respectively, students commented:

To assist other students who struggle with sight, it is a very beneficial project that means all can learn equally and with ease without feeling like an outcast in any way. (P4 - self)

The feeling that they may be able to help people who are less fortunate than them. (P21 - others)

Seeing how their participation would be useful to disabled students, was thought by six students to be a motivator. On the other hand, only one student thought that would motivate other students. On this motivating factor in relation to themselves, one student said:

I would try it to see how it effects visually impaired students. If I can see that my descriptions are helping them to understand, then I would describe more images (P15)

The same student said in relation to what would motivate others:

Similar to what would motivate me but also the possibility of prize(s). (P15)

Only two students thought that being in competition with other students would motivate them to participate to the project, whereas six students thought that would motivate their peers. One of the students who was motivated by this factor said:

Instead of asking me to describe ask two people to describe the picture to each other…..they can compete against other teams of two or against a high score (P5)

Another participant said about what would motivate other students:

Students can compete against each other and motivate each other to participate. (P12)

The time required to participate and the flexibility of participation was a concern for two students, who thought that participation should not interfere with their studies or the study of other students. A participant said:

If the times were flexible, so that I could participate without it interfering with my studies. (P1)

Another participant said about what would motivate other students:

depending on work/time required. (P7)

Only one participant thought that improving the CV would motivate him or other students to participate, he said:

A reference for my CV (P7)

Table 5.7 Mean and median ratings, standard deviation and One-Sample Wilcoxon Signed Rank tests for the 12 motivational factors

Motivational Factors	Mean	Median	SD	W *p*
Your sense of altruism, wanting to help other students (Altruism)	5.7	6.0	1.2	3.98 0.00
Getting academic credits (Academic credits)	5.2	6.0	2.0	2.45 0.014
Enhancing your job opportunities in the future (job opportunities)	5.2	6.0	2.0	2.50 0.012
Being paid for your efforts (Money)	5.0	5.0	1.9	2.42 0.015
The fun and entertainment of the activity (Fun/entertainment)	4.7	5.0	1.4	2.21 0.027
Being connected with other students on your course (Being connected)	4.7	4.5	1.3	2.39 0.017
Improving your academic skills (Improving skills)	4.7	5.0	2.2	1.58 n.s
Knowing that you are contributing to a large project (Contributing)	4.6	5.0	1.7	1.78 n.s
Drawing attention to your skills (Attention to skills)	4.0	4.0	1.7	-0.71 n.s
Being in a competition with other students (Competition)	3.8	4.0	2.2	-0.96 n.s
To pass the time (To pass time)	3.0	3.0	1.7	-2.59 0.010
The social recognition you would receive (Social recognition)	2.8	2.5	1.6	-2.92 0.004

Another student thought that getting notification when his descriptions being used by lecturers would motivate him to participate, he said:

If I was notified each time my description was actually used in a lecture, I would feel like I was making more of a difference (P6)

Interestingly, the partially blind student thought that the idea of helping those who are in a worse situation to himself was a very motivating factor.

I am a partially blind cs [Computer Science] student and even though [I] have some vision would like to help others who are in a worse situation than me (P25)

In the second part of the questionnaire students were asked to rate to what extent each of 12 motivational factors would motivate them to participate, using 7 point Likert items (1= Not at all to 7=Very much). Overall, students rated altruism as the highest motivational factor to participate in the image description project. On the other hand, social recognition was the lowest rated motivational factor. The results are illustrated in Table 5.7 and Figure 5.5.

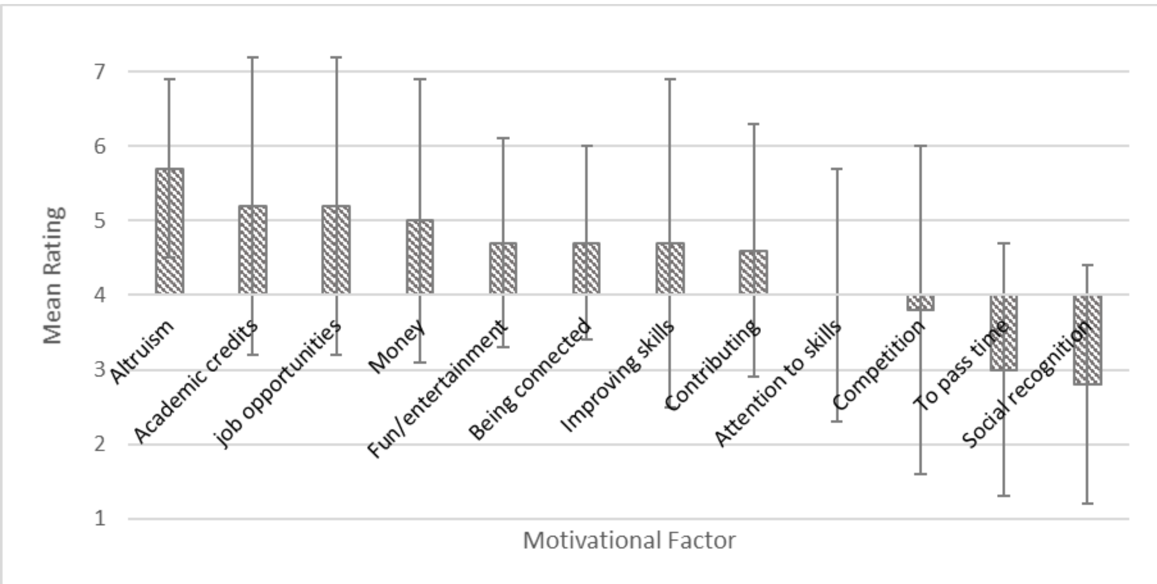


Figure 5.5 Mean rating and standard deviation for the 12 factors as motivations for self against the neutral point on the rating scale

The following sections present more detailed results for each factor, examples of students' ratings and comments about each factor can be found in Appendix C.2.4 .The motivational factors are presented from the most important to the least important according to Table 5.7.

Students thought they would be highly motivated by their sense of altruism and wanting to help disabled students. Table C.13 shows some of the students' comments about this factor.

Getting academic credits as result of participating in the DescribeIT project was very appealing to students (see Table C.14). However, as some of the students commented this sounded nice but not necessarily possible to apply.

Students thought they would be motivated by the chance of enhancing their job opportunities in the future, as shown in Table C.15. However, those who were not motivated by this factor could not see how participation in crowdsourcing project could be mirrored in enhancing their job opportunities.

Earning extra cash (see Table C.16) was thought to be a motivation to participate in the DescriebIT project. However, in their comments students mentioned the amount of money they could earn out of their participation, hence to motivate students by money to participate in socially responsible crowdsourcing, one must think about the right amount of money to pay per task.

Overall, students thought they would be motivated by the fun and entertainment of the image description task. Table C.17 shows some of students' thoughts about this factor.

The idea of being connected with other students in the course through the DescribeIT project was thought by students to be an important factor (see Table C.18). However, students who were not in favour of this factor thought that they can be connected with their peers by other means.

Students were neutral in their feeling about improving their academic skills. While some students could not see how the project would improve their academic skills (see Table C.19), others, thought the idea of trying to describe images from their learning materials could potentially help improving their academic skills.

The scale of the project was not a key factor for students to participate in the DescribeIT project, for as shown in Table C.20 some thought that meant that their contribution would be very little and others thought that meant that the work could be done very quickly.

Students were also neutral about their feeling about drawing attention to their skills through participating in the project. Table C.21 illustrate some students' comments.

Students were neutral in their feeling about being in competition to describe images for blind and partially sighted students (See Table C.22).

Students were not at all motivated by the idea of passing the time, students thought they are already have too much work to do (See Table C.23)

As shown in Table C.24, students thought that although social recognition would be a nice result of their participation, this would not be a reason for them to participate in the DescribeIT project.

Phase 2: Behaviour in the Crowdsourcing Project

None of the 26 students who participated in Phase 1 or any of the other students who were invited to participate in Phase 2 did so. This was in spite of being sent reminder emails, and their indications of their likelihood of participating when responding to the Phase 1 questionnaire. Three students in the InMot condition accessed the project and viewed five images. Only one student in the ExMot condition accessed the project and viewed two images. None of these students provided any descriptions.

This surprising outcome led me to extend the study into a follow-up phase to attempt to understand their lack of participation in the crowdsourcing project.

5.4.4 Phase 3: Exploring students' lack of participation

5.4.4.1 Introduction

The purpose of this additional phase of the study was to attempt to understand why this group of students did not participate in Phase 2 of the study. As the participation in Phase 1 and 2 of the study had been so low, it was decided to use a personal interview with a monetary incentive for participation to elicit a good response rate and as much information as possible from students.

5.4.4.2 Method

5.4.4.2.1 Design

Semi-structured interviews were conducted to investigate why none of the students who had participated in Phase 1 of the study (the online questionnaire) had not participated in the crowdsourcing project (Phase 2) of the study, despite the fact that in Phase 1 these students reported that they were motivated to participate in a crowdsourcing project to support their blind and partially sighted peers. In addition, at the end of the interview session data were collected about students' general altruistic behaviour and their interaction with disabled people, using the SRA scale (Rushton et.al. 1981) and the IDP scale (Gething, 1994) respectively. The aim was to investigate if this sample of non-participating students were particularly low in their altruism and attitudes towards people with disabilities.

5.4.4.2.2 Participants

The same group of 169 students who had been invited to participate in Phases 1 and 2, were invited to participate, and a total of 19 students responded (a response rate of 11.2%) comprising 5 women and 14 men. Their ages ranged from 18 to 21 years old, with a mean age of 18.8 years (SD=1.0). This group comprised 7 students who had been assigned to the InMot condition in Phase 2 (22.5% response rate), including 3 women and 4 men and 12 students (8.6% response rate) including 2 women and 10 men who had been assigned to the ExMot condition. The age range for students from the InMot condition was from 18 to 19 years old, with a mean age of 18.1 years (SD=0.4). The age range for students from the ExMot condition was from 18 to 21 years old, with a mean age of 19.3 years (SD=0.9).

The majority of students were from the United Kingdom. All 7 students who took part from the InMot condition were from the United Kingdom. Whereas of the 12 students who took part from the ExMot condition, 9 were from the United Kingdom, two were from China and one from Germany.

All students reported using social media, such as Facebook and Twitter. However, in the InMot condition only two students indicated that they had participated in crowdsourcing projects at least once in the past. None of ExMot condition students reported participating in any crowdsourcing projects before. Students from both conditions said they have not been actively seeking to participate in crowdsourcing projects, because they either never came cross anything which attracted them to participate or they thought they did not have the technical skills necessary to participate.

As an incentive to participate in the interview, students were offered a £15 Amazon gift voucher for students of the InMot condition and £10 Amazon gift voucher for students of the ExMot condition. (More details about why students in the two conditions were offered a different amount of money is explained in the Procedure section, below).

5.4.4.2.3 Materials

A semi-structured interview schedule covered the following topics:

1. A general introduction about the DescribeIT project and its importance to blind and partially sighted students

2. Students' thoughts on the results of the self-reports of motivations (Phase 1)
3. Willingness to help blind and partially sighted peers
4. Participation in the image description project (Phase 2 of the study)
5. Willingness to participate in the DescribeIT project if they were recruited again
6. Demographic data and online activities.

Students also completed the SRA Scale (Rushton et al., 1981) to collect information about their self-reported general altruistic behaviour and the IDP Scale (Gething, 1994) to collect information about their interaction with disabled people. For more details about the scoring of both scales see Chapter 2, section 2.8.

5.4.4.2.4 Procedure

A recruitment email was sent out to the group of students who had been invited to Phase 1 and Phase 2 (InMot condition) to participate in the study, with a doodle URL to sign up for an interview at a convenient date. A £5 Amazon gift voucher was offered as an incentive to take part in an interview of approximately 15 minutes. Unfortunately, after five days of sending this recruitment email, no-one had responded. Therefore, a reminder email was sent, but again no one responded. A lecturer who was teaching this group of students at the time was asked to invite students personally to participate. Again, no students offered to participate. Finally, a further recruitment email with an offer of a £15 Amazon gift voucher for participation was sent. In response to this email, seven students were recruited.

Another recruitment email was also sent out to the group of students who had been invited to Phase 1 and Phase 2 (ExMot condition) to participate in the study, with a doodle URL to sign up for an interview at a convenient date. To increase the chances of recruitment, a £10 Amazon gift voucher was offered as an incentive for participation instead of £5. Six students responded to this recruitment email and further six to a reminder email sent five days later.

At the start of the session, students were given an information sheet to read about the study (see Appendix C.2.5), any questions they had were answered and then they signed a consent form.

In the first part of the interview students from both motivation conditions were given a summary of the project and its importance, including a summary of the aim of each of the previous phases of the project. Students then were asked the questions scheduled for the interview. At the end of the interview session students were thanked for their time and valuable input to the research, they also were debriefed about the study.

5.4.4.3 Results

Students were presented with the results of Phase 1, and a copy of Figure 5.5 was shown to them, students were then asked what their thoughts about the results, and whether they thought it is a true representation of what would motivate them and their peers to participate in the DescribeIT project. The results showed that most students agreed with the Phase 1 results, two students did not agree. One of these latter students thought that the job opportunities and fun/entertainment motivational factors should probably be less important to students; and the other of these students thought that the motivational factors money and improving skills should probably be more important to students. In addition, it worth noting that one participant with some psychology background as she said stated that:

I believe 'Altruism' should be further down! People would love to think they are altruistically motivated but the truth they are not

However, she continued:

this is just because of my background otherwise I think this is a good representation of what would motivate students.

When students were asked whether they remembered the request to participate in the DescribeIT project, 6 (31.6%) students out of the 19 said they had in fact participated. One student said that he had started the task, but was worried that his descriptions were not adequate, so had not submitted any descriptions, he said:

I was not happy with my description to click on submit …I was thinking what if it is not good enough and it went through

When he was asked whether the guidelines were helpful or not he said*:*

I do not remember reading them, because I thought it was clear what is expected from me.

However, for the other students it was not clear whether they did not want to admit they had not participated, were genuinely misremembering, or that they had also started the image description task but had not submitted anything. Comments from several students indicated that they had looked at the project, for example one student stated:

… some of the images were challenging it is hard to put images into words briefly…it was more challenging than I thought it would be … (P4).

Another student said*:*

Some images were easy to describe others were more difficult" he continued *"I thought that the lay out of the project was good and easy to understand and navigate.*

This suggests that students may have found the task difficult and daunting and perhaps felt it would be more time consuming than had been intended. For example, some students commented on how much time they would be prepared to dedicate to the project:

It depends, if it [needed a] huge commitment no, if it [is] a case of few minutes here and there that is OK … (P6)

It depend[s] on how much time…15 to 30 minutes…why not?

The other 13 (68.4%) students did say they had not participated and gave a number of obvious reasons for this (See Table 5.8), including not receiving the emails, seeing the emails but not getting around to participating, being involved with other research projects and being busy with their university work, and in particular having deadlines for assignments.

Table 5.8 Students' reasons for not participating in the DescribeIT project

The reasons behind the lack of participation	**N, %**	**Typical comments**
Being busy with their university work, and in particular having deadlines for assignments	5 (38%)	*I was busy with assignments.*
Not receiving the emails	3 (23%)	*Can't remember receiving the email.*
Seeing the emails but not getting around to participating	2 (15%)	*I did not had time…. got the recruitment email, but was busy… unfortunately*.
Being involved with other research projects	2 (15%)	*I was involved with different experiments…with other two studies*.
Cannot remember why	1 (8%)	*can't remember why … maybe stuff at home*.

None of the students knew a blind or partially sighted person. The majority of students stated that they would be happy to participate in the project if they were invited again. In addition, nine students thought that talking to the researcher about the project made them see how this project was important to visually disabled students.

To investigate whether this sample of non-participating students were particularly low in altruism and attitudes to people with disabilities, their Self-Report Altruism Scale (SRA) and Interaction with Disabled Persons (IDP) scores were examined.

The mean SRA score was 40.5 (SD = 6.2). This mean was compared to previous published results for the SRA. Rushton et al. (1981) provided results for the SRA from 611 Canadian university students. While these university students were in a different country, and the data quite old, this was the best comparison data set which could be found. The comparison was somewhat complicated by the fact that four questions from the original SRA were not used, as they were not thought to be suitable for the groups of students (see section 5.3.3.1 for further details). However, when the original SRA scores provided by Rushton et al. (1981) were scaled and averaged, a mean of 44.5 was obtained. A one-sample Wilcoxon Signed Rank test found that the scores for the current group of students was significantly lower than the median of SRA scores provided by Rushton et al. (1981) ($W=-2.6$, $p<0.05$).

The mean IDP score was 68.1 (SD= 10.0). Comparison with published was easier in this case as Brown et al. (2009) reported results on the IDP for British undergraduate students. Their mean IDP score was 62.31. A one-sample Wilcoxon Signed Rank test found that the scores for the current group of students was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=2.3$, $p<0.05$).

5.4.4.4 Discussion

In light of the results of the follow up Phase, it was decided to invite the group of students who had been invited to take part in the previous phases (1,2, and 3) to participate in the crowdsourcing project again to make a further investigation of their willingness to participate. There were several reasons for this decision. Firstly, it was because the majority of the students interviewed said they would participate if they were sent another invitation to participate in the project. Secondly, all of them indicated that they were not as busy with studies or assignments at the time of the interviews as they had been during Phase 2 of the study. Lastly, nine students indicated that talking to the researcher about the project in the interview session increased their understanding of the importance of the project.

5.4.5 Phase 2, Round 2 of Behaviour in the Crowdsourcing Project

5.4.5.1 Introduction

Although the timing of Study 3 was chosen very carefully to make sure that students were not overwhelmed by assignments or exams at the time to increase participation, the results of Phase 3 suggested that it is quite difficult to achieve that with full-time students. In addition, the University of York has a considerable research community and students are invited to take part in many research studies, and offered money as compensation for their time and effort in most of these studies. However, motivated by the results of Phase 3, it was decided to investigate students' willingness to participate to the DescribeIT project again.

5.4.5.2 Method

One week after the end of the interview sessions two recruitment emails were sent out. A recruitment email was sent out to the original group of 31 students in the InMot condition using the same email used in Phase 2 round 1 [See Appendix C.2.2] to participate in the study. A separate recruitment email was sent out to the 138 ExMot condition students using the same email used in Phase 2 round 1 [See Appendix C.2.1] to participate in the study. One reminder email was sent out a week later.

5.4.5.3 Results

After two weeks, none of the students in the InMot condition had participated or accessed the project. This included the six students who participated in the interviews and who said they would participate if they were invited to the project after the interview.

However, seven students (5% of 138 invited) from the ExMot condition accessed the DescribeIT project. Overall these students viewed 63 slides, a mean of 9.0 (SD=7.8) slides per student. However, only four students provided image descriptions, 17 descriptions in total. The number of images described range from 1 to 12, a mean of 4.3 (SD=5.3) descriptions per student. The other three students only viewed the images, without providing any descriptions. The number of images viewed range from 1 to 4, with a mean of 2.7 (SD=1.5)

5.4.6 Discussion and Conclusions

This study investigated what undergraduate students thought would motivate them to participate in the DescribeIT project and compared their self-reports with their actual behaviour in participating in such a project. The study investigated two motivational factors, an intrinsic factor (the usefulness of the project to blind and partially sighted students) and an extrinsic factor (increasing the students' of understanding the teaching materials). For results summary see Tables 5.12 and 5.13 end of this chapter.

Overall, students showed an interest in participating in the DescribeIT project. Students rated the likelihood of their participation and the participation of other students significantly above the neutral, with a significantly higher likelihood rating of participation to themselves in comparison to their peers. The latter finding is similar to the findings of the previous studies (Studies 1 and 2), which suggested the effect of the fundamental attribution error. As explained in Study 1 (See Discussion section 4.4) the fundamental attribution error proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits. Hence, when students were asked to predict their own participation, they thought of the specific situation of helping disabled students, but when they were asked to predict the participation of other students, they thought of the general helpfulness of other students, not in the context of helping disabled students.

The self-report of motivations shows that students thought they would be motivated by a mix of intrinsic and extrinsic motivations. However, as the nature of DescribeIT project under investigation in this programme of research might affect students' tendency to respond in a socially desirable way, students were asked what they think would motivate them and others to participate to the project, to reduce socially desirability bias.

The overall results suggest that students gave a socially desirable response to both Likert items ratings and the open-ended question of what they thought would motivate them. For example, when students were asked what would motivate them to participate, ten students thought that financial rewards would motivate them, in comparison to almost double the number (19 students) who thought that would motivate their peers. Furthermore, the results of the open-ended question when students were asked about what would motivate them do not match their ratings to the same motivational factor in the second section of the online questionnaire. In the former (open-ended question) financial rewards were mentioned most often and in the later (rating item) although financial reward was significantly above neutral it was rated as the fourth most important motivational factor for them to participate. Nonetheless, overall students were motivated by money. The finding of this study is in agreement with the findings of Studies 1, and 2 (although the rating of money was not significantly different from midpoint of the scale, but the overall results suggest that student were motivated by money). These results also agree with those from studies by Lakhani et al. (2007) and Brabham (2008, 2010), which all reported the importance of monetary rewards as a self-reported motivation.

Similar to Studies 1 and 2, students in the present study thought that they would be highly motivated by their sense of altruism, and wanting to help other students, which is in line with the results of self-reports of motivation in previous research as well (e.g. Kuznetsov 2006, Oreg and Nov 2008). However, one needs to consider three findings: Firstly, that in Phase 1 of the present study, seven students (27%) thought that altruism would motivate them to participate. Secondly, there total lack of participation in Phase 2 round 1 of the present study, and the very low participation rate in round 2. Thirdly the very low participation rate in the InMot condition (Altruism) of Phase 2 in study 2. These three findings suggest that the results of the present study and Study 2 students overreported the importance of this factor to participate to the DescribeIT project in comparison to their actual behaviour.

Although getting academic credits was not suggested by students when they were asked in the open-ended question about what would motivate them and others, students' ratings of this motivational factor showed that students were significantly motivated by this factor. This result is in line with the results of Study 1, but not with Study 2 in which participates were neutral about this factor.

Enhancing one's job opportunities in the future was also a highly-rated factor, as students thought participating in a project like this looks good in their CVs and would increase employability in the competitive job market. Yet only one participant suggested the important of improving one's CV for future career when they were asked in the open-ended question what would motivate them or others. The findings of Study 1 on this factor agrees with the findings of the present study. However, in Study 2 students were neutral about this factor, mainly because they were not motivated to pursue a career in a related field.

The image description task was not perceived by students of Studies 1 and 2 to be a fun task to do, hence they were neutral about their feeling of the important of this factor to participate to the DescribeIT project. However, in the present study students thought that they would be motivated by the fun of the image descriptions task, as students thought that "enjoyment" would be important motivate them to do the image description task. This finding is aligned with the findings of other research (Brabham 2008, 2012, Hossain 2012, Kaufmann, et al. 2011) which reported the importance of the fun factor.

Previous research (Von Ahn and Dabbish, 2004; Brabham, 2012) highlighted the importance of the community in crowdsourcing projects and being connected to other members of the crowd in such projects. This particular finding is similar to the finding of the present study. However, it contrasts with the results of Studies 1 and 2; in Study 1 students were not motivated by this factor at all, and in Study 2 students were neutral about their feeling about this factor.

The results about improving one's skills have been consistent throughout the studies (Studies 1, 2, and the present study), the results of all studies indicate that students' ratings were neutral. However, this result is not similar to those from Brabham's studies (2008, 2010, 2012), in which he reported the importance of this factor. Nevertheless, improving or developing artistic skills at iStockphoto (2008) or Threadless (2010) may not be similar to improving students' academic skills, being more personal rather than career-oriented skills.

Similarly, but to a lesser degree, students were neutral about drawing attention to their skills. This sounds at odds compared to the results about enhancing job opportunities (which was significantly important), and is also contrary to the finding of Brabham's (2008) study, in which iStockphoto members reported the importance of their participation in the project to improve their marketability and getting a better job.

Interestingly, when students were asked about what would motivate them and other students to participate in the open-ended question, two students thought that competition would motivate them, in comparison to six students who thought this would motivate other students. In addition, two students suggested that being in competition with other students would not only motivate them to participate but also would increase the quality of their descriptions. However, overall students were neutral in their ratings about the competition factor, which contrasted with the findings of the first two studies in the programme of research in which students rated this motivational factor significantly low. While one would expect students to be competitive, especially as they have to be competitive in their search for jobs when they graduate, yet being competitive in helping others does not seem to be appealing to them.

Analogously with the previous studies (Studies 1 and 2), passing the time and the social recognition students would receive were rated significantly low. The results of passing the time factor is understandable, due to the fact that students of the present study and also the previous ones (Studies 1 and 2) were full-time students and therefore do not have much spare time to worry about passing. Comparable results were reported by Brabham (2008) with the iStockphoto members, they were not motivated by this factor. Lastly, the results on social recognition factor in the previous studies (Studies 1 and 2) agree with the results of the present study in which students did not think that the idea of getting recognized for helping disabled students appealing. This result contrasts to Hossain's (2012) literature review in which he cited the importance of this factor by several studies.

The fact that no students participated in the crowdsourcing project in Phase 2 was totally unexpected, given the results of the questionnaire in Phase 1 which clearly showed that students thought they would participate in the DescribeIT project. The interviews with a small number of the students showed a range of sensible reasons, including pressure of studying, simple forgetfulness and finding the task daunting. However, this result highlights the limitations of relying on self-report measures when studying motivation; self-report measures may not be accurate predictors of participants' actual behaviour. In the case of socially responsible topics, self-report measures may be particularly vulnerable to social desirability biases (Fisher 1993) – few people may want to say that they are unwilling to help their disabled peers, even in an anonymous questionnaire. In the case of this crowdsourcing project, there may be several particular factors which may have led to the non-participation.

Firstly, as hinted at by one of the students interviewed in Phase 3, creating good image descriptions is not a trivial task. Students may well not have realized that until they actually looked at the project. Although the number of students who accessed the project in the first round of Phase 2 was very low (only four students), in the second round it was slightly higher, with seven of the students, all in the ExMot condition accessing the project, but then only four students provided any descriptions, and not many descriptions at that. These results do suggest that students may have found the task too difficult when they looked at the project. I will explore this possibility in Studies 8 and 9, in which I will compare participation rates in a crowdsourcing project which asks students to describe images (as in this study) with a project which asks students to do a much simpler task, to tag individual objects within the image. The tagging task could still be useful in providing descriptions of images for visually disabled students, as lecturers or students could build up on them to create useful descriptions. In addition, in Study 10 I will investigate how students perceive the level of difficulty of the image description task before describing any images and how their perception of the level of difficulty changes when they have described some images.

Secondly, although the instructions in the DescribeIT project emphasized that the descriptions would benefit blind and partially sighted students taking the course in the future, as far as the students know there were no blind or partially sighted students in the courses the students were taking and they may have felt it was unlikely that blind or partially sighted students would take the course in the future. This may have made students feel the tasks was not a very meaningful or useful one on which to spend time. This is something which will be investigated in Study 4 which will be run with the students of the Department of History, at the University of York, a department which has had a number of blind and partially sighted students in the last few years and is known for its excellent record of supporting blind and partially sighted students.

Finally, the intrinsic and extrinsic motivational factors chosen for the two conditions in Phase 2 were ones which emerged from the open-ended questions of the previous Studies 1, 2 and Phase 1 of the present study. Thus, they had been reported a number of times as motivators by students. However, they did not actually motivate this group of students, especially in the case of the ExMot condition, in which no one explicitly reported the importance of this factor.

In conclusion, Phase 1 of the study showed that students were mainly motivated by altruism and money. Seeing how their participation is helpful to blind and partially sighted students is a motivator emerged from students' answers. Phase 2 of the study shown, as Antin and Shaw (2011) highlighted, that relying on self-reports of motivations in crowdsourcing studies, as numerous researchers have, may not be accurate in terms of predicting behaviour.

5.5 STUDY 4: The Effect of a Positive Environment on Participation in a Socially Responsible Crowdsourcing Project

5.5.1 Introduction

In Studies 2 and 3 students were recruited from the Departments of Computer Science and Theatre, Film and Television, both are not known for accepting blind and partially sighted students. Thus, students of those studies may have felt that it was meaningless to spend their time describing images, as there were no blind or partially sighted students who could use the descriptions. Therefore, it was decided to conduct a study with students of the Department of History at the University of York, a department is known for its excellent record of accepting blind and partially sighted students. However, this did not mean necessarily that potential individual students would personally know blind or partially sighted students, but they might well be aware of the positive environment in their department for students with visual disabilities and have a greater awareness of the importance of making digital learning resources accessible to such students. In general lecturers in the Department of History were enthusiastic about the DescribeIT project, and the project was presented by one of them to the potential students as a very useful tool for blind and partially sighted students of the Department of History.

5.5.2 Method

5.5.2.1 Design

The study investigated how having a positive environment for blind and partially sighted students in the department could affect participation in a socially responsible crowdsourcing project to support these students. As with Studies 2 and 3, this study consisted of two phases, as follows:

Phase 1: Investigated how students perceive different motivational factors that would affect their participation and the participation of other students in the DescribeIT project. The same online questionnaire used in Studies 2 and 3 was used again [Appendix B.1.3].

Phase 2: Explored the behaviour of the students while using the DescribeIT project. There was only one condition, in which students were presented with the same video used in the InMot condition of Study 3, in which a blind history student talked about his daily challenges with inaccessible learning materials and how the DescribeIT project could help him and other blind and partially sighted students.

The hypothesis was that students in this department will show a higher participation rate than students in the previous two studies (Studies 2 and 3), because they have a greater expectation that their participation will impact blind and partially sighted peers.

5.5.2.2 Participants

236 History students were invited to participate in the study. 36 (15.3%) students participated in Phase 1, one student participated in Phase 2. The 36 students in Phase 1 comprised 14 men and 22 women. Their age range was from 19 to 23 years old, with a mean age of 19.9 years (SD=0.89). Most students were from the UK, except for three who were from China, Hungary and Ireland. All students reported using social media. However, only two students reported that they had participated in crowdsourcing before. One participant said he participated in video games project because it *gave power to the consumer.* The other participant said that she participated in "Zooniverse[8] - Operation War Diary" because it was *fun and interesting*.

As an incentive to complete the questionnaire, students were entered into a prize draw for one of 10 £10 Amazon vouchers.

5.5.2.3Materials

Phase 1 used the same online questionnaire as in Studies 1 to 3 (see Appendix B.1.3). For Phase 2 used the scales as Studies 2 and 3 (see Appendix C.1.3).

5.5.2.4 DescribeIT project

The DescribeIT project used in this study was the same one discussed in Chapter 3, section 2.3. To help students create good image descriptions, the project information page provided guidelines on how to describe images for blind and partially sighted people and an example description of a typical image.

The DescribeIT Project presented students with a PowerPoint pack from a seminar in one of the courses students were taking at the time of the study, "Using Visual Materials in Historical Research". Figure 5.6 presents the layout of the project. Students were told that they could describe as many or as few images as they wished over as many sessions as they wished.

[8] https://www.zooniverse.org/

Figure 5.6 The DescribeIT project layout for Study 4

5.5.2.5 Pilot Study

The DescribeIT project for this study was piloted by one Computer Science PhD student at the University of York. No technical problems were detected. The participant did not take part in the main study.

5.5.2.6 Procedure

My supervisor Professor Helen Petrie visited one of lectures for the cohort of history students. She was introduced by the lecturer to the students. In his introduction, the lecturer emphasized the important of the DescribeIT project to the Department of History. Professor Petrie then talked to the students about the aim of the research and its importance to blind and partially sighted students.

A recruitment email was sent out to students on the same day as this lecture (see Appendix C.3.1) inviting them to take part in Phase 1 of the study. Students were told specifically both in the lecture and in the recruitment email that: *The University of York is trying to be particularly welcoming to blind and partially sighted students, and the Department of History has an excellent record of accepting blind and partially sighted student, so we thought plotting this idea in History would be a good idea*.

A reminder email was sent a week later.

Two weeks later students were invited to take part to the DescribeIT project (Phase 2) (see Appendix C.3.2). Two reminder emails were sent out a week apart.

5.5.2.7 Data Analysis

Data for some of the dependent variables were not normally distributed, thus nonparametric tests were carried out for all analyses. Open-ended questions were analysed with content analysis as discussed in Study 1 (See 4.2.5 for more details).

5.5.3 Results

5.5.3.1 Phase 1: Self-Reports of Motivation

Students were asked to rate their likelihood to participate in the DescribeIT project and the likelihood of other students participating, on 7 point Likert items (1=Not at all, 7=Very likely). Students gave a mean rating of 5.1 (SD=1.3) that they would participate in the project. However, students thought that likelihood of their peers to participate in the project less than them, with a mean rating of 4.2 (SD=1.4). Figure 5.7 shows the pattern of mean ratings for these two questions. To investigate if the likelihood of students' ratings of the likelihood of participation was significantly above the midpoint of the rating scale, one-sample Wilcoxon signed rank tests were conducted comparing the ratings with the midpoint rating of 4. The results are summarised in Table 5.9. This showed that students' ratings of their own likelihood to participate in the project were significantly higher than the neutral midpoint ($W=4.0$, $p < .001$). However, their ratings for their estimation of the likelihood that other students would participate were not significantly different from the neutral midpoint of the rating scale. A Wilcoxon signed-rank test showed that students rated the likelihood of themselves participating in the project significant higher than the likelihood of their peers participating in the project ($Z = -2.9$, $p < 0.005$).

Ten motivational factors emerged from the open-ended question which asked students about what would motivate them to participate in the DescribeIT project, including: wanting to help others, rewards, knowing a blind person, more information about the project, time required, making learning more accessible, university recognition, convenient location, regular reminders and talking to them about the project before their lecture. Table 5.10 summarizes these motivational factors which students thought would motivate them and other students to participate in the DescribeIT project.

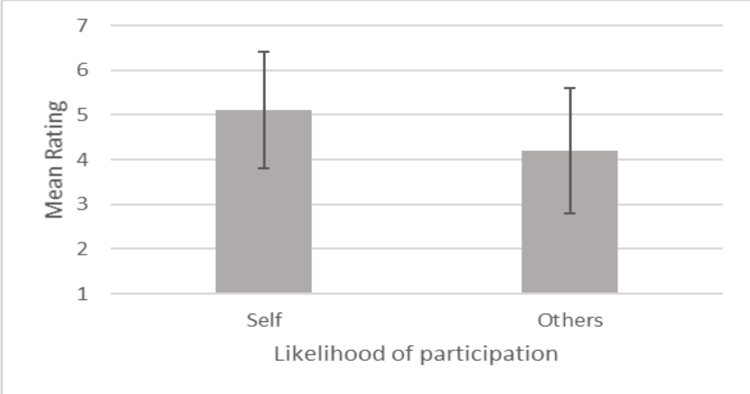


Figure 5.7 Mean ratings and standard deviation on the likelihood of own and other students' participation in the DescribeIT project

Table 5.9 Mean and median ratings and standard deviation on the likelihood of own and other students' participation in the DescribeIT project

Likelihood of participation	**Mean**	**Median**	**SD**	**W**	***p***
Self	5.1	5	1.3	4.0	0.000
Others	4.2	4	1.4	1.1	0.252

For many students (about 33%) the desire to help blind and partially sighted students was a key motivator. The following is some of the students' thoughts on that:

Desire to assist in the development of provisions for partially sighted students (P5)

Helping other students - especially as there are a lot of images involved in history lectures/seminars (P11)

The fact that I could help people that are genuinely suffering (P17)

In addition, making the learning materials more accessible was an important factor. A participant said:

To aid other students in any way possible, making Uni [University] more accessible for all (P36)

Helping those who are visually impaired to enjoy the same benefits and use the same sources to enrich their degree as someone who doesn't suffer from the same issue (P31)

The chance of receiving rewards such as Amazon vouchers, cash or free food was thought to be a strong motivator by students, nine students (25%) thought it would motivate them, in comparison to 14 students (39%) who thought it would motivate other students:

A reward such as Amazon vouchers or cash (P23)

Being paid or recognition in some form (P34)

Table 5.10 Categorization of factors which students thought would motivate themselves and other students to participate in the DescribeIT project (from open-ended question)

Motivational factors	**Self** (N, %)	**Others** (N, %)
Wanting to help others	12 (33.3%)	11 (30.5%)
Rewards (e.g. vouchers, cash, food)	9 (25%)	14 (39%)
Knowing a blind person	5 (14%)	1 (3%)
More information about the project (advertising before they start and reports of progress after they start)	4 (11%)	6 (17%)
The time required to participate not too long	4 (11%)	3 (8.3%)
Making learning more accessible	4 (11%)	1 (3%)
Convenient location	2 (6%)	2 (6%)
University recognition	1 (3%)	4 (11%)
Regular reminders	1 (3%)	1 (3%)
Professor Petrie talking to them before lecture	1 (3%)	0 (0%)

However, other type of rewards like an award from the university (e.g. credit toward the York Award[9], a scheme run by the University to encourage students to take on socially responsible extra-curricular activities) was less important especially when students talked about themselves in comparison to other students. For example, one student said:

Some form of incentive-counts towards York award (P19)

Five students mentioned that they know someone who is a blind or partially sighted, a factor they said would motivate them to participate to any socially responsible project to support people in need:

One of my housemates is partially sighted so I am aware of the difficulties faced by partially sighted students when it comes to module material (P2)

[9] https://www.york.ac.uk/students/work-volunteering-careers/skills/york-award/

My grandma was blind, so I understand the importance of projects such as these because I'm aware of what they meant to her when she was alive (P4)

I know a quite nice chap in my seminar who is partially sited [sighted] (P6)

I have a partially sighted friend and know the limitations that this creates (P26)

A student said:

I am mildly dyslexic and have a very disabled brother, and think it would be fantastic to have resources like this so more disabled people could participate (P32)

While this student did not know specifically someone who is blind or partially sighted, her disability and that of her brother is a motivator to her to help others.

However, only one student speculated about how knowing a blind or partially sighted person would motivate other students. The student said:

Knowing that they have friends who are partially sited [sighted] in their course so that they can put a face to the people they are helping. (P6)

Four students (11%) noted that getting information before participation and after participation would motivate them to participate, in comparison to six (17%) who thought this would motivate others to participate. Students thought that advertising the project to potential students and providing them with the necessary information would motivate them to participate. In addition, students thought that seeing the impact of their participation on disabled students would be a motivator as well. Some comments from students on these issues included:

Further information about the project goals (P7)

Lots of publicity and reaffirming the reason behind the project (P14)

Really seeing that it's actually being used to help people (P36)

Interestingly, a student said that: *Helen [Petrie] speaking about the project before my lecture (P10)* was a motivator to makes her participate to the project.

Although the proposed DescribeIT project is an online project, which allows students to describe as many or as few images as they wish at any time they wish, two students (6%) talked about the importance of the location and four students (11%) talked about the time required to participate in the project as limitations on their ability and motivation to participate:

Not require a large amount of time that would take away from my studies (P16)

I would be motivated to participate in the project if it was relatively quick and simple, rather than taking a long period of time. If it was something you could maybe complete in 15 minutes or less that would be great (P21).

It was also suggested by one participant that it would be good to do the description task during seminars: *if we were asked to do it during seminars that would probably offer some motivation (P14)* whereas, another participant suggested some sort of gathering to do the task: *If it only took place around one day a week and for a doable amount of time in a destination easy to get to (like on campus) (P27).*

Lastly, a participant thought getting regular reminders would increase the likelihood that he and other students would participate: *Regular reminders to do so (P35).*

In the second part of the questionnaire students were asked to rate to what extent each of 12 motivational factors would motivate them to participate, using 7 point Likert items (1= Not at all to 7=Very much). Results are summarized in Figure 5.8 and Table 5.11. Overall, students rated their altruism as the highest motivational factor to participate in the image description project. Whereas, social recognition was the lowest motivational factor.

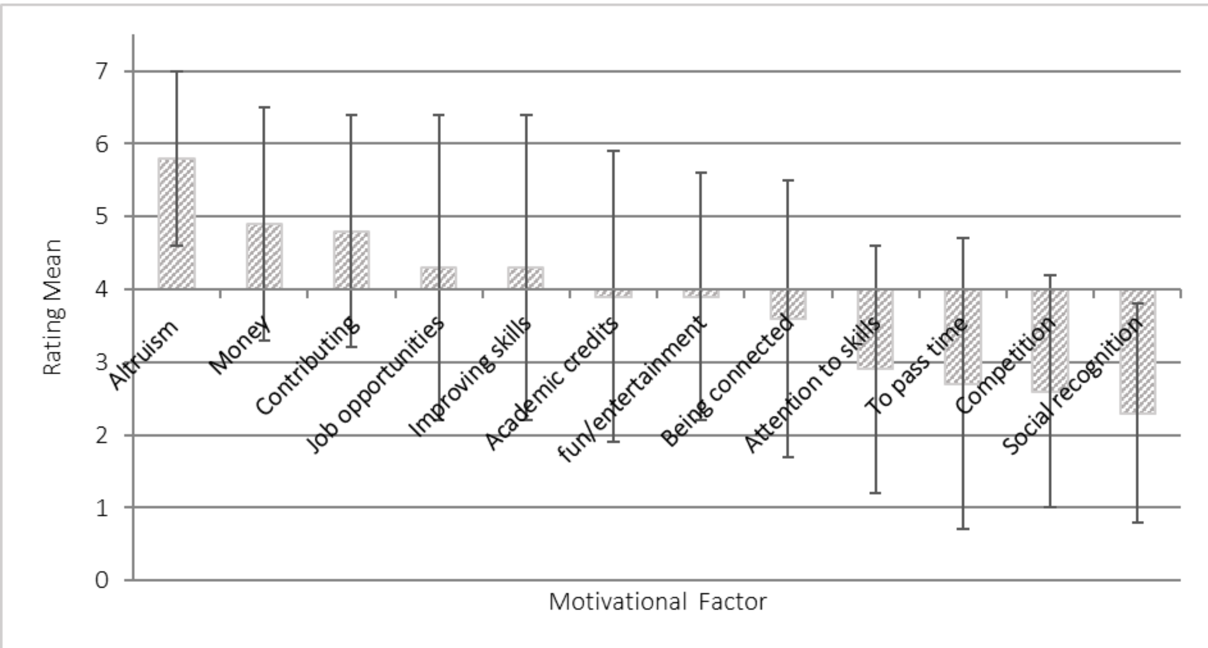


Figure 5.8 Mean ratings and standard deviation for the 12 factors as motivations for self against the neutral point on the rating scale

Table 5.11 Mean, standard deviation and One-Sample Wilcoxon Signed Rank tests for the 12 motivational factors

Motivational Factors	**Mean**	**Median**	**SD**	**W** ***p***
Your sense of altruism, wanting to help other students (Altruism)	5.8	6.0	1.2	4.8 0.000
Being paid for your efforts (Money)	4.9	5.0	1.6	2.9 0.004
Knowing that you are contributing to a large project (Contributing)	4.8	5.0	1.6	2.6 0.010
Enhancing your job opportunities in the future (Job opportunities)	4.3	4.5	2.1	0.5 n.s.[1]
Improving your academic skills (Improving skills)	4.3	5.0	2.1	0.9 n.s.
Getting academic credits (Academic credits)	3.9	4.0	2.0	-0.3 n.s.
The fun and entertainment of the activity (fun/entertainment)	3.9	4.0	1.7	-0.4 n.s.
Being connected with other students on your course (Being connected)	3.6	4.0	1.9	-1.4 n.s.
Drawing attention to your skills (Attention to skills)	2.9	2.0	1.7	-3.3 0.001
To pass the time (To pass time)	2.7	2.0	2.0	-3.5 0.001
Being in a competition with other students (Competition)	2.6	2.0	1.6	-3.9 0.000
The social recognition you would receive (Social recognition)	2.3	2.0	1.5	-4.5 0.000

1 n.s = Not significant

A series of one sample Wilcoxon signed rank tests were carried out to see if students' motivations were statistically different from the neutral midpoint rating of 4 (See Table 5.11). The results indicated that students' ratings were significantly higher than the neutral midpoint rating for the following factors: altruism (p<0.0001), money (p<0.005), and contributing (p<0.01). Whereas, students' ratings were significantly below the midpoint for: To pass time (p<0.001), competition (p<0.0001), and social recognition (p<0.0001).

Students had the option to explain their responses on the Likert items for each of the 12 motivational factors. Appendix C.3.4 shows some of the students' comments.

5.5.3.2 Phase 2: Actual Behaviour in Crowdsourcing Project – DescribeIT

Only one student participated in the DescribeIT project, she produced five descriptions. As result no statistical analysis was made of the Phase 2 questionnaire.

5.5.4 Phase 2 Round 2: Actual Behaviour in Crowdsourcing Project – DescribeIT

5.5.4.1 Introduction

The results of Phase 1 were promising, and it was hoped that students would participate in Phase 2 of the study. However, unfortunately only one student participated to the DescribeIT project, in spite of two reminder emails being sent to students. To make sure that the lack of participation was not due to any of the reasons reported by students in Study 3, interview sessions (See section 5.4.4 for more details), it was decided to run Phase 2 of the study again, but this time only with the 36 students who participated in Phase 1 of the study.

5.5.4.2 Method

Four weeks after the closure of the first round of Phase 2, a recruitment email was sent out to the 36 students who had participated in Phase 1 of the study, inviting them to participate in the DescribeIT project. A reminder email was sent one week later.

5.5.4.3 Results

Unfortunately, similarly to the result of the first round of Phase 2, only one student participated in the DescribeIT project, describing only 6 images. It was not possible to check whether it was the same student who participated in both occasions, as in both occasions it was an anonymous participation to the DescribeIT project and in the second occasion the student did not take part in the online questionnare.

5.5.5 Discussion and conclusions

This study aimed to investigate what students in the Department of History, a department with a supportive environment for visually disabled students, thought would motivate them to participate in the DescribeIT project. The Department of History was particularly chosen because of its excellent record of accepting blind and partially sighted students at University of York, which the researcher had hoped would improve students' participation rate in the DescribeIT project. For results summary see Tables 5.12 and 5.13 end of this chapter.

The study showed that students reported an interest in participating to the DescribeIT project. Students rated the likelihood of their participation to the DescribeIT project significantly above the neutral. However, their ratings to the participation of other students was neutral. Students likelihood of participation ratings for themselves were significantly higher than the ratings of participation to their peers. The fact that students rated the likelihood of themselves participating higher than other students is similar to the findings of the previous studies (Studies 1, 2 and 3), which suggests the effect of the fundamental attribution error. As explained in Study 1 (See 4.4) the fundamental attribution error proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits. Hence, when students were asked to predict their own participation, they thought of the specific situation of helping disabled students, but when they were asked to predict the participation of other students, they thought of the general helpfulness of other students, not in the context of helping disabled students.

To reduce the tendency of social desirability bias in self-report, students were asked about what would motivate them, and also what would motivate others to participate in the DescribeIT project in an open-ended question. The answers showed a mix of intrinsic and extrinsic motivational factors. Similar to the previous studies (Studies 1, 2, and 3) wanting to help others and financial rewards where the leading motivational factors. However, in contrast to all the previous studies (Studies 1, 2, and 3) students in the present study thought that financial rewards were less important to themselves and other students than the motivational factor of wanting to help others. Yet, they thought that wanting to help others is slightly more important to themselves (12 students) than to other students (11 students), whereas, money was more important to other students (14 students) than to themselves (9 students). This finding was similar to those of Studies 1, 2, and 3, in which students put more emphasis on intrinsic motivational factors when they wrote about themselves in comparison to when they wrote about other students, whereas, they put more emphasis on extrinsic motivational factors when they wrote about other students in comparison to when they wrote about themselves. In addition, knowing a blind or partially sighted person was thought by some students to be an important motivational factor for participation in the DescribeIT project. On the other hand, only one student thought that could be a motivational factor for other students. Getting more information about the project before participation (e.g. why the project is important?) and after participation (e.g. on the progress of the project) was also thought to be an important motivational factor by some students.

The results of the open-ended questions (when students were asked about what would motivate them) and the results of their ratings were matched (i.e. when the same motivational factor that students named was one of the 12 motivational factors they were asked to rate). For example, the results of the open-ended question showed the importance of the motivational factor, altruism and the same motivational factor was rated the highest among the 12 motivational factors. Similarly, financial rewards were thought by students to be the second most important factor in the open-ended question, and this matched the results of the rating of the same motivational factor, which was also significantly above neutral. Furthermore, the results of these two factors (altruism and money) matches the results of previous research, for example Kuznetsov (2006), and Oreg and Nov (2008) reported the importance of the altruism factor, and Lakhani et al. (2007) and Brabham's (2008, 2010) importance of financial rewards factor.

For students knowing that they are contributing to a large project was thought to be important factor for participation, as their rating for this factor was significantly above neutral, which is in agreement with the results of Study 1. Perhaps this is why students of the present study wrote about the importance of getting information about the project before participating, including advertisements about the project and progress reports after participating, when they were asked about what would motivate them and other students to participate, as these requirements can be expected from large projects.

The chance of enhancing one's job opportunities in the future was rated as neutral by students, those who were motivated by this factor thought participating in this project would look good on their CV, whereas, others who were not motivated by this factor principally gave two reasons for their ratings: the first, they would like the idea of enhancing their job opportunity as result of their participation, but this would not be the motivator to initiate participation. The second, they were not motivated to pursue a career in a field related to the project. The results of this factor were similar to the results of Study 2, but contrasted to the results of Studies 1 and 3, in which students reported this factor as being important.

Consistent with the results of all the previous studies (Studies 1 ,2, and 3) the students in the present study thought that they were neutral about the motivator of improving one's skills. However, this contrasted with the results of Brabham's studies (2008, 2010, 2012), in which he reported this factor as being important. Nevertheless, as explained previously, improving or developing artistic skills at iStockphoto (2008) or Threadless (2010) may not be similar to improving students' academic skills, being more personal than career-oriented skills.

Students were neutral in their feelings about the idea of getting academic credits as result of participating to the DescribeIT project. This result is in line with the results of Study 2, but not with the results of Studies 1, and 3, in which students thought they were significantly motivated by this factor.

The findings of the present study showed that students were neutral about the fun and entertainment of the activity, which is in agreement with the results of Studies 1 and 2, but contrary to Study 3 and some previous research (Brabham 2008, 2012, Hossain 2012, Kaufmann, et al. 2011). These studies found that the fun and entertainment was the leading self-reported intrinsic motivation factor. As discussed in Study 1, this could be due to the nature of the description task, as students did not think describing an image was a fun task to do, whereas for example Brabham's studies examined artistic design tasks.

Students were neutral in their thoughts about the motivator of being connected with other students in their course, which is in line with the findings of Study 2. However, the results of Study 3 showed that students thought they would be motivated by this factor, and similar to previous research (e.g. Von Ahn and Dabbish, 2004; Brabham, 2012) which has highlighted the importance of the community and being connected within crowdsourcing.

In Studies 1 and 2 and in the present study students thought that they would not be motivated by drawing attention to their skills. This finding contrasted with the results from Brabham's (2008) study, in which he reported the importance of drawing attention to members' artistic skills at iStockphoto. In line with the results of Studies 1, 2, and 3 and previous research (Brabham 2008), students in the present study thought that they would not be motivated by the motivational factor of passing the time. This is understandable since all students were full-time university students, and probably they have a very busy schedule of study-related work.

Lastly, students thought that they would not be motivated by being in competition with other students nor by receiving social recognition. The finding on the competition factor is similar to those of Studies 1 and 2, in which students did not find it appealing to compete in relation to helping disabled students. In contrast to results from Hossain (2012) on the importance of social recognition, the results of the present study and Studies 1, 2, and 3 showed that students thought that they would not be motivated by getting social recognition for their participation.

In Phase 2, only one student participated in the DescribeIT project, even after sending the recruitment email and two reminder emails. This result was surprising for two reasons. Firstly, students were told that this would be particularly helpful to blind and partially sighted students in the Department of History, and they were presented with an interview with a blind History student talking about how would this project would be helpful to him (so they knew that there were blind and partially sighted students at the Department, even if they did not know them personally). In addition, the blind student talked about the importance of students' participation, and how they understand the context of the images more than anyone else. Secondly, the results of Phase 1 of the present study showed that students were motivated to participate to the DescribeIT project. As a result, the students who participated in Phase 1 were invited again four weeks after the first round of data collection to participate to the DescribeIT project, with one reminder email. Yet again only one student participated. It was not possible to check whether it was the same student who participated in both occasions, as the student chose to participate anonymously in both round 1 and 2 to the DescribeIT project. The student in round 1 participated in the online questionnaire, but not in round 2.

A comparison between the participation rate in Studies 2, 3, and 4, showed that students in the present study had the lowest participation rate (0.8%), a slightly higher participation rate was noted in Study 3 (4%) in the second round, and the highest participation rate was noted in Study 2 (11%), with the MSc. students.

In conclusion, the self-reported motivation showed that students were mainly motivated by their sense of altruism and money. Other motivational factors such as knowing that they contribute to a large project and knowing a blind or partially sighted person thought to be a driving factors to participate. Unfortunately, the positive environment in the department of History did not influence students to participate to the DescribeIT project. This study, similar to Studies 2 and 3, has shown as Antin and

Shaw (2011) highlighted, that relying only on self-reports of motivations in crowdsourcing studies may not be accurate in terms of understanding people motivation.

Table 5.12 Summary of the results of Studies 1 to 4

		Study 1	Study 2	Study 3	Study 4
Number of participants (phase 1)		271 students	27 students	26 students	36 students
Number of participants (phase 2)		N/A	8 students	0 phase 3: 19 students	2 students
Number of descriptions		N/A	86 descriptions	17 descriptions (round 2)	11 descriptions
Likelihood of participations		Results of One-sample Wilcoxon tests			
Self		+**	n.s	+**	+**
Others		n.s	n.s	+*	n.s
Motivational Factors					
Altruism		+**	+**	+**	+**
Money		+**	n.s	+*	+**
Academic credits		+**	n.s	+*	n.s
Job opportunities		+**	n.s	+*	n.s
Contributing		+**	n.s	n.s	+*
Improving skills		n.s	n.s	n.s	n.s
Fun/entertainment		n.s	n.s	+*	n.s
Being connected		-*	n.s	+*	n.s
Attention		-**	-*	n.s	-**
Competition		-**	-*	n.s	-**
Social recognition		-**	-*	-**	-**
Pass time		-**	-**	-*	-**
SIMS Scale	InMot Condition				
Intrinsic Motivation	InMot Condition	N/A	-	-	-
Identified Regulation	InMot Condition	N/A	-	-	-
External Regulation	InMot Condition	N/A	-	-	-
Amotivation	InMot Condition	N/A	-	-	-
SIMS Scale	ExMot Condition				
Intrinsic Motivation	ExMot Condition	N/A	-	-	N/A
Identified Regulation	ExMot Condition	N/A	-	-	N/A
External Regulation	ExMot Condition	N/A	-	-	N/A
Amotivation	ExMot Condition	N/A	-	-	N/A

Table 5.13 Correlation between number of images and SIMS, IDP, and SRA scale of Studies 1 to 4

SIMS Scale	Study 1	Study 2	Study 3	Study 4
Intrinsic Motivation	N/A	-	-	-
Identified Regulation	N/A	-	-	-
External Regulation	N/A	-	-	-
Amotivation	N/A	-	-	-
IDP Scale				
Sympathy	N/A	-	-	-
Discomfort	N/A	-	-	-
SRA Scale	N/A	-	-	-

Chapter 6

The motivation of Arab students to participate in a socially responsible crowdsourcing project

6.1 Introduction

A key strength of crowdsourcing is that it can attract a crowd from many geographical locations and cultural contexts. This can be beneficial to crowdsourcers (those seeking to have a task done, See Chapter 2, section 2.2.1), who can access a large pool of potential crowd members at any time of the day. However, it may also pose a challenge for them to understand how to motivate a potentially geographically and culturally diverse crowd (those who undertake the tasks, See Chapter 2, section 2.2.1). Theories of cultural psychology (e.g. Hofstede 1983, 2001) have suggested that people's motivations may vary between cultures, since culture strongly influences our values, attitudes, and behaviours. For example, the Arab countries have different cultural values in comparison to North American and European countries. In the Arab countries people have a strong attachment to their communities. In Arab and Islamic culture, it is very important to give charity and help less fortunate people (Singer 2013), hence Arab people tend to be involved in volunteer work. This could be due to one of the cultural dimensions discussed by Hofstede, individualism/collectivism. Hofstede (1983) defines individualism as a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families. Its opposite, collectivism, he defines as representing a preference for a tightly-knit framework in society in which individuals can expect their relatives or members of a particular in-group to look after them in exchange for unquestioning loyalty.

Drawing on Hofstede's theory, I decided to investigate the attitudes and motivations of students from two Arab countries, Libya and Saudi Arabia, in relation to crowdsourcing to help people with disabilities. In his research[10], Hofstede found that Libyans and Saudis had very collectivist values, whereas British people (the majority of participants in the studies reported so far) had amongst the highest individualist values in the

[10] https://geert-hofstede.com/countries.htm

world. Therefore, the hypothesis for this study is that students from Arab countries will display more collectivistic values than students from the UK. This means that Arab students will be both more likely to say that they will participate in a crowdsourcing project to help disabled people and more likely to actually participate in such a crowdsourcing project than British students.

This Chapter will present Study 5 which investigated Arab students' perception of what would motivate them to participate to the DescribeIT project and their actual participation. The study will investigate how different intrinsic and extrinsic motivational factors can affect the participation of Arab students in the DescribeIT project. It will also continue to investigate the magnitude of the social desirability bias in reporting motivations to participate in socially responsible crowdsourcing projects. Students' situational motivation (refers to the motivation of an individual while engaging with the image description task, see Chapter 2, section 2.8) was measured to assess the constructs of intrinsic motivation, identify regulation, external regulation and amotivation (Guay, Vallerand, and Blanchard, 2000). The hypothesis is that students who are intrinsically motivated will produce more image descriptions in the DescribeIT project than those who are extrinsically motivated. It was also hypothesized that students' attitudes toward disabled people would affect their behaviour in the DescribeIT project. Therefore, students' attitudes towards disabled people were measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994). Finally, it was hypothesized that students' level of altruism would affect their behaviour in the DescribeIT project. Thus, students' altruistic behaviour was measured using the Self-Report Altruism Scale (SRA) (Rushton et.al. 1981).

6.2 Method

6.2.1 Design

The study investigated how intrinsic and extrinsic motivational factors can affect the participation of Arab university students in a socially responsible crowdsourcing project. As in the previous studies in this research programme, the study consisted of two phases:

Phase 1*:* investigated how Arab university students perceive different motivational factors that would affect their participation and the participation of other students in a crowdsourcing project to describe images for their blind and partially sighted peers. This used the same online questionnaire as used in Studies 1 to 4 (Appendix B.1.3).

Phase 2: investigated the behaviour of Arab university students in a crowdsourcing project to describe images for their blind and partially sighted peers. There were two conditions: one with an intrinsic motivation and one with an extrinsic motivation. The instructions in the intrinsic motivation condition (InMot) was about the feeling of satisfaction one gets when helping others, while in the extrinsic condition (ExMot) the instructions were about how participation could potentially improve students' skills. These particular intrinsic and extrinsic motivational factors were chosen based on the results of Phase 1 of the present study, both were rated significantly above neutral (InMot, rating mean= 6.2; ExMot, rating mean= 4.7, see section 6.3.1).

In Phase 2 of the study, the dependent variables included the number of students who participated in the study, the number of attempts students made to describe images and the quality of the image descriptions produced.

Other dependent variables were:

- Students' motivation while doing the tasks, measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000).
- Students' overall attitude toward disabled people, measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994).
- Students' self-reported level of altruism, measured using Self-Report Altruism scale (SRA) (Rushton et.al. 1981).

6.2.2 Participants

Approximately 300 students were invited to take part in this study. A total of 92 (approximately 30% response rate) university students from Libya and Saudi Arabia participated in Phase 1 of the study, comprising 68 women and 16 men (eight students did not answer the gender question). Their age range was 18 to 47 years old, with a mean age of 27.9 (SD = 6.6) years old. Students were undergraduate (36, 39%) and postgraduate students (56, 61%) from range of disciplines, including Computer Science, Education, Engineering, Medicine, and Pharmacy at Omar Al-Mokhtar University and Benghazi University in Libya and King Saud University in Saudi Arabia.

All the students reported using social media such as Facebook. However, only two of the students reported having participated in crowdsourcing projects, and did not specify which projects these were.

As an incentive to participate in the study, students were entered into a prize draw for one of 10 Amazon vouchers equivalent of GBP 10.

In Phase 2 of the study 89 (approximately 30% response rate) students participated in the DescribeIT project, 44 students in the InMot condition and 45 students in the ExMot condition. A total of 67 (approximately 75% response rate from those who participated in Phase 2) students participated in the Phase 2 questionnaire, 37 students in the InMot condition and 30 students in the ExMot condition.

As an incentive to complete the Phase 2 questionnaire, participants were entered into a prize draw for one of 10 Amazon vouchers each worth equivalent of GBP 10.

6.2.3 Materials

Phase 1 of the study used the same online questionnaire used in previous studies 2, 3 and 4 (see Appendix B.1.3). This comprised four sections: an explanation of the project; questions about students' perceived likelihood of their own participation in the project and that of other students; questions about what factors would motivate them to participate in the project; and demographic and background information. For more details about each section see Chapter 4, section 4.2.3. As students were doing their undergraduate and postgraduate studies in English it was not necessary to translate the questionnaire into Arabic.

The Phase 2 online questionnaire was the same one as used in studies 3 and 4, and comprised of four sections: Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000), Interaction with Disabled Persons Scale (IDP) (Gething, 1994), Self-Report Altruism Scale (SRA) (Rushton et.al. 1981), and demographic questions. For more details about the scales see Chapter 2, section 2.8.

6.2.4 DescribeIT project

The DescribeIT project used in this study was discussed in Chapter 3, section 2.3. To help students create good image descriptions, the project information page provided guidelines on how to describe images for blind and partially sighted people and an example description of a typical image from the PowerPoint slides to be described. These were provided both in Arabic and English. The Arabic version was translated by having a PhD student, from the Department of Education at the University of York, who speaks both Arabic and English fluently translate the English version into Arabic. The Arabic translation was then given to a Lecturer of Arabic literature at the University of Benghazi - Libya to translate back into English, without having access to the original English version.

The PowerPoint slides used in this study were from an introductory course on computer science at the fourth level of the primary school in Libya. Similar content is used in the Saudi Arabian school curriculum. The project used a PowerPoint presentation containing 15 slides from the course. However, only 13 slides contained images, the aim was for each image to be described 25 times on each of the conditions, which meant a total of 325 descriptions on each condition.

Once students started the project and read the instructions, they were presented with a slide and a text-box in which to type their description of the image on the slide. Figure 6.1 illustrates the project layout. Students were advised they could provide their descriptions either in Arabic or English. This was to minimize any effect on the quality of descriptions of having to work in a second language.

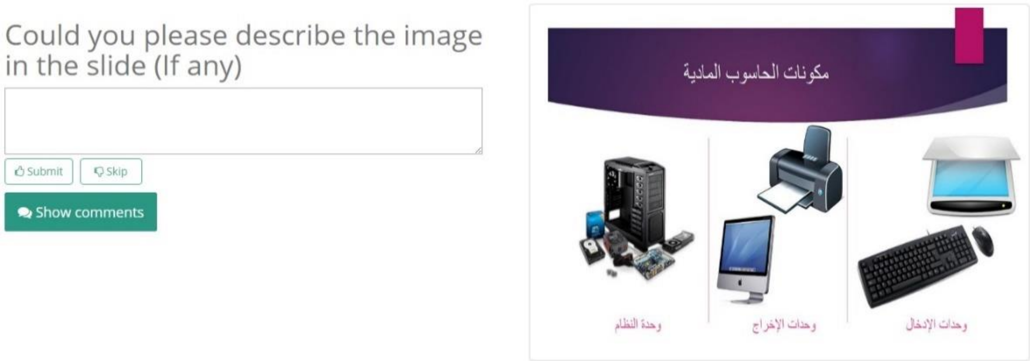


Figure 6.1 The DescribeIT Project Layout for Study 4

6.2.5 Data Analysis: Image description quality

To assess the quality of the descriptions, the number of words in the image descriptions was used as a simple metric for the quality of the description. Typically, the longer a description, the more detail it has and the more useful it will be to a visually disabled student. In addition, the relation between the number of objects mentioned in a description and the number of words of the same description was measured for a random sample of 130 descriptions to investigate this as a potentially more accurate metric.

6.2.6 Pilot Study

Three students from the University of York (two research students in Computer Science and one research student in Education) piloted the project, all were Arabic speakers. The pilot study revealed some technical issues with the Arabic descriptions, as the project did not render Arabic letters correctly. This problem was resolved. These students did not take part in the main study.

6.2.7 Procedure

Students were recruited through university contacts (i.e. lecturers) in Libya and Saudi Arabia. A recruitment email was sent out to undergraduate and postgraduate students at Omar Al-Mokhtar University and Benghazi University in Libya and King Saud University in Saudi Arabia, in a range of humanities and science departments, including Computer Science, Education, Languages, Medicine, and Pharmacy. In Phase 2 students were randomly assigned into one of the two conditions. Two reminder emails at approximately five days' intervals were sent out.

For Phase 1 the online questionnaire was available for one week, and for Phase 2 the DescribeIT project was available for four weeks along with the Phase 2 questionnaire.

6.2.8 Data Analysis

Some variables in data collected were not normally distributed. Thus, it was decided to use nonparametric tests for all analyses.

Qualitative analysis of the two open-ended questions about students' motivations to take part in the DescribeIT and open-ended questions about other students' motivations to take part in the DescribeIT were carried out using content analysis (see section 4.2.5 for more details).

6.3 Results

6.3.1 Phase 1: Self-Reports of Motivation

In Phase 1 of the study students rated the likelihood that they would participate in the DescribeIT project and the likelihood that other students on their course would participate in the project, both on 7 point Likert items (1 =Not at all likely to 7 = Very likely). Figure 6.2 shows the mean ratings for these two questions. Overall, students gave a mean likelihood rating that they would participate of 5.5 (SD = 1.9), and a mean rating that other students would participate of 4.3 (SD =2.0). To investigate if the likelihood of students' ratings of the likelihood of participation was significantly above the midpoint of the rating scale, one-sample Wilcoxon signed rank tests were conducted comparing the ratings with the midpoint rating of 4 (See Table 6.1). This showed that students' ratings of their own likelihood to participate in the project were significantly higher than the neutral midpoint (W=5.6, N=92, $p< 0.001$). However, their ratings for their estimation of the likelihood that other students would participate were not significantly different from the neutral midpoint of the rating scale (W =1.4). A Wilcoxon signed-rank test also showed that participants rated the likelihood of themselves participating in the project significantly higher than from the likelihood of their peers participating in the project ($Z = -5.0$, $N=92$, $p< 0.001$).

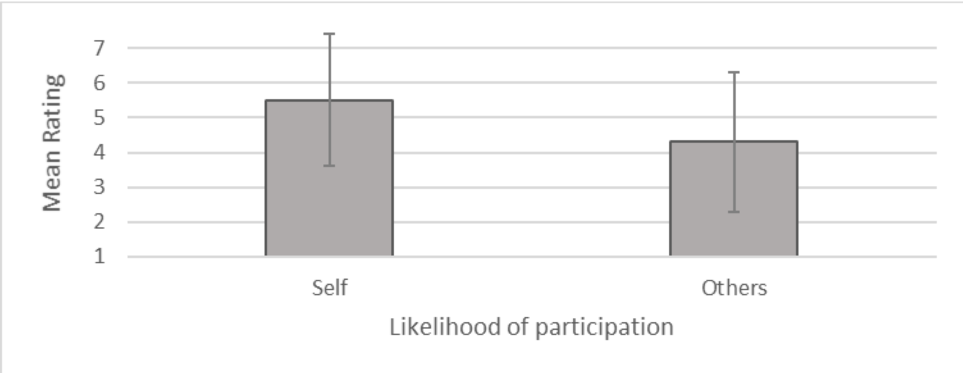


Figure 6.2 Mean ratings and standard deviations of likelihood of participating in the DescribeIT project, for self and other students

Table 6.1 Mean and median ratings, standard deviations (SD) and the results of one-sample Wilcoxon signed rank tests of students' rating of their own and others likelihood of participation.

Likelihood of participation	**Mean**	**Median**	**SD**	**W**	***p***
Self	5.5	6.0	1.9	5.6	0.000
Others	4.3	4.0	2.0	1.4	n.s

In an open-ended question, students noted a range of intrinsic and extrinsic factors that would motivate them and other students to participate in the DescribeIT project, these are summarised in Table 6.2.

The desire to help disabled students was the most frequently mentioned motivational factor that students thought would motivate them and other students to participate to the DescribeIT project. A student said:

To help people with special needs who need this kind of facility (P16)

Another student said:

Help blind and partially sighted students to learn (P70)

Table 6.2 Categorization of factors which students thought would motivate themselves and other students to participate in the DescribeIT project (from open-ended question)

Motivational factors	**Self** (N, %)	**Others** (N, %)
Wanting to help disabled students	33 (35.9%)	23 (25%)
Supporting research	20 (21.7%)	7 (7.6%)
Improving the education of disabled people	7 (7.6%)	4(4.3%)
Being involved in an interesting project	7 (7.6%)	5 (5.4%)
Improving Skills	3 (3.3%)	2 (2.3%)
Getting more information about disabled people	3 (3.3%)	5 (5.4%)
Monetary rewards	2 (2.3%)	9 (9.8%)
Knowing that the project is useful for disabled students	2 (2.3%)	5 (5.4%)
Knowing a blind or partially sighted student	2 (2.3%)	1 (1.1%)
Receiving academic credits	0 (0%)	2 (2.3%)

The second most frequently mentioned motivational factor was willingness to support research, especially research that would improve the life of disabled students, one student said:

I am interested in research concerning disabilities in my country (P19)

Other students said:

Helping in science research (P23)

Support research (P50)

The idea of improving the education of disabled people was thought by seven students to be an important factor to them, but only four thought that would motivate their peers. Some students said:

For equality of opportunities and access to knowledge (P18)

providing a better chance for blind people to learn (P42)

Being involved in an interesting project was thought by seven students to be a motivational factor, whereas only five thought that would motivate other students, a student said:

I think it is a useful one. I was touched by the idea of having challenged people understand what others see which is something we normally take for granted, and never think of those who do not see what we do. (P69)

Improving skills and knowing a blind or partially sighted person were less important to students, as only three thought they would be motivated by the Improving skills factor, and only two thought it would motivate others. Whereas two student thought knowing a blind person would motivate them in comparison to only one who thought it would motivate other students. Some students' comments about these factors include:

My desire to learn (P30) [Skills]

Knowing a friend can benefit from the project. (P24) [knowing a blind or partially sighted person]

Students thought the following motivational factors would be more important to other students than to themselves: getting more information about disabled people, monetary rewards, knowing that the project is useful for disabled students, and getting academic credits. The following are some examples of students' comments about these factors:

More awareness of the problem. Knowing that there are blind people who are suffering would indeed raise the level of awareness among people. (P53) (getting more information about disabled people)

Financial rewards (P69) (Monetary rewards)

If they know this project would benifit [benefit] those people with sight disability [visually impaired] (P10) (knowing that the project is useful for disabled students)

Giving them marks!! (P1) (Academic credits)

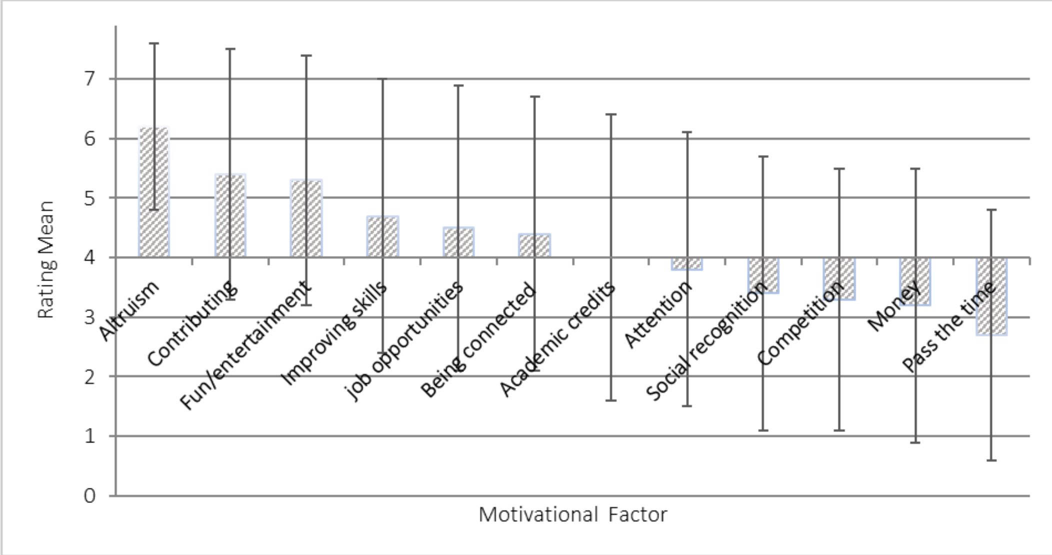


Figure 6.3 Mean ratings and standard deviations for the 12 factors as motivations for self to participate in the DescribeIT project against the neutral point on the rating scale

Table 6.3 Mean and median ratings and standard deviations for the 12 rated motivational factors with results of one-sample Wilcoxon signed rank tests

Motivational	**Mean**	**Median**	**SD**	**W**	***P***
Altruism	6.2	7.0	1.4	7.53	0.000
Contributing	5.4	6.0	2.1	4.89	0.000
Fun/entertainment	5.3	6.0	2.1	4.36	0.000
Improving skills	4.7	5.0	2.3	2.44	0.015
job opportunities	4.5	5.0	2.4	1.56	n.s.
Being connected	4.4	5.0	2.3	1.32	n.s.
Academic credits	4.0	4.0	2.4	-0.200	n.s.
Attention	3.8	4.0	2.3	-1.29	n.s.
Social recognition	3.4	3.0	2.3	-2.520	0.012
Competition	3.3	3.0	2.2	-2.84	0.004
Money	3.2	3.0	2.3	-3.44	0.001
Pass the time	2.7	2.0	2.1	-5.01	0.000

In the second part of the questionnaire students were asked to rate to what extent each of 12 motivational factors would motivate them to participate, using 7 point Likert items (1= Not at all to 7=Very much). Figure 6.3 illustrates the students' ratings for the 12 motivational factors. Overall, students rated their sense of altruism as the highest rated motivational factor to participate in the DescribeIT project, whereas to pass time was the lowest rated motivational factor.

A series of one-sample Wilcoxon signed rank tests were carried out to see if students' motivations were statistically different from the neutral midpoint rating of 4 (See Table 6.3). The results indicated that the motivational factors of Altruism, Contributing, Fun/Entertain and Academic skills were rated significantly above the midpoint, whereas Social recognition, Money and Pass time were rated significantly below the midpoint.

6.3.2 Phase 2: Students' Behaviour in Crowdsourcing Project

Overall 89 students participated in the DescribeIT project, 44 in the InMot condition (29% of those invited) and 45 in the ExMot condition (30%). A total of 584 descriptions were created in the project, a mean of 8.7 descriptions per student (SD=3.9), being: 238 in the InMot condition, a mean of 9.4 descriptions per student (SD=3.9), and 346 in the ExMot condition, a mean of 7.9 descriptions per student (SD=3.8). A Mann-Whitney test was conducted to see if there was a significant difference between the number of images produced in the InMot condition and the number of images produced in the ExMot condition, this showed no significant difference.

Table 6.4 The Frequency and Percentage of the Word Count Ranges in the Image Descriptions

Words Range	InMot		ExMot		Overall	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
1 to 25 words	165	69%	269	78%	434	74%
26 to 50 words	46	19%	65	19%	111	19%
51 to 75 words	21	9%	8	2%	29	5%
76 to 100+ words	6	3%	4	1%	10	2%
Total	**238**	**100%**	**346**	**100%**	**584**	**100%**

Table 6.4 illustrates the frequency and percentage of the word count range in the image descriptions. Overall the mean number of words per description was 20.6 (SD= 17.6, range 2 to 112 words), the mean number of words per description in InMot condition was 22.7(SD= 20.8, range 2 to 112 words), the mean number of words per description in ExMot condition was 19.2 (SD= 14.8, range 2 to 101 words). To investigate if there was a significant difference in the number of words between the two conditions, a Mann-Whitney test was conducted. This showed no significant difference in the

number of words between the two conditions. A Spearman correlation was carried out to investigate the relationship between the word count and the number of objects mentioned in a description, which showed a significant positive correlation. $r_s(130) = 0.7$, $p<0.001$.

For the Phase 2 questionnaire 67 students participated, 37 students in the InMot condition and 30 students in the ExMot condition. The mean and median ratings and standard deviations for the InMot and ExMot condition scores on the four sub-scales of the SIMS are summarized in Table 6.5 and 6.6 respectively. To investigate whether students' scores on the SIMS sub-scales were statistically different from the neutral midpoint of 4, a series of one-sample Wilcoxon signed rank tests were conducted for the students in the InMot and ExMot conditions separately. The results of these tests are also summarized in Tables 6.5 and 6.6 respectively. For both the InMot and ExMot conditions, Intrinsic Motivation scores were significantly higher than the midpoint ($p<0.001$), whereas scores for External Regulation ($p<0.05$) and Amotivation ($p<0.001$) were significantly lower than the midpoint.

Table 6.5 Means, median, standard deviations and summary of significance tests for SIMS subscales in the InMot condition

SIMS Sub-scale	**InMot (N=37)**				
	Mean	**Median**	**SD**	**W**	***p***
Intrinsic Motivation	4.8	5.0	1.4	3.1	0.002
Identified Regulation	3.9	3.5	1.9	-0.5	n.s
External Regulation	3.3	3.3	1.2	-2.8	0.005
Amotivation	2.4	2.5	1.0	-5.2	0.000

Table 6.6 Means, median, standard deviations and summary of significance tests for SIMS subscales in the ExMot condition

SIMS Sub-scale	**ExMot (N=30)**				
	Mean	**Median**	**SD**	**W**	***P***
Intrinsic Motivation	5.1	5.1	1.5	3.3	0.001
Identified Regulation	4.3	4.3	1.6	1.1	n.s
External Regulation	3.4	3.5	1.3	-2.0	0.042
Amotivation	1.7	1.3	0.9	-4.7	0.000

On the Self-Report Altruism Scale (SRA) students in the InMot condition scored a mean of 46.6 (SD=8.6, range: 27 - 64), whereas students in the ExMot condition scored a mean of 49.1 (SD=9.6, range: 26 – 66). As shown in Table 6.7 a Mann-Whitney test showed no significant difference between the two conditions. A one-sample Wilcoxon Signed Rank test found that the scores for the Arab students was significantly higher than the median of SRA scores provided by Rushton et al. (1981) ($W=-2.8$, $p<0.01$). This mean that this group of students was more altruistic than those reported in Rushton et al. (1981).

Table 6.7 Means, standard deviations and summary of significance tests for SIMS, IDP, SRA scales in the two conditions

Measures	**InMot (N=37)**		**ExMot (N=30)**		**(sig)**
SIMS Sub-scale	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**	
Intrinsic Motivation	4.8 (1.4)	5.0	5.1 (1.5)	5.1	n.s
Identified Regulation	3.9 (1.9)	3.5	4.3 (1.6)	4.3	n.s
External Regulation	3.3 (1.2)	3.3	3.4 (1.3)	3.5	n.s
Amotivation	2.4(1.0)	2.5	1.7 (0.9)	1.3	0.004
IDP Sub-scales					
Sympathy	25.1 (3.5)	25.0	25.0 (3.9)	26.0	n.s
Discomfort	15.5 (5.3)	16.0	14.1 (5.7)	15.5	n.s
SRA Scale	46.6 (8.6)	48.0	49.1 (9.6)	49.5	n.s

On the Interaction with Disabled Persons Scale (IDP) students the InMot condition scored a mean of 15.5 (SD=5.3) on the Discomfort sub-scale, and a mean of 25.1 (SD=3.5) on the Sympathy sub-scale. Whereas, in the ExMot condition results showed a mean of 14.1 (SD=5.7) for the discomfort sub-scale, and 25.0 (SD=3.9) for the sympathy toward disabled people sub-scales. A series of Mann-Whitney tests were conducted to investigate if there were differences on the overall IDP score and IDP sub-scales between the InMot and ExMot conditions. This shown no significant differences between the two conditions. A one-sample Wilcoxon Signed Rank test found that the scores for the Arab students was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=6.7$, $p<0.000$). This mean that Arab students were more dis-comfortable around disabled students than British undergraduate students reported in Brown et al. (2009) study. This could be due to the fact that disabled people are not fully integrated in the societies in Arab countries. Hence, people do not know how to interact with them.

Another series of Mann-Whitney tests were conducted to investigate if there were differences on the SIMS sub-scales between the InMot and ExMot conditions. The results are summarized in Table 6.7, which shows that there was only one significant difference between the two conditions, which was on the SIMS Amotivation sub-scale.

To investigate whether there was a relationship between the number of images produced and the scores on the SIMS, IDP sub-scales and the SRA, a series of Spearman correlations were conducted for students in the InMot and ExMot conditions separately. These are summarized in Table 6.8, which shows that there were significant correlations only for the ExMot condition. In this condition, there was a significant positive correlation between the number of images described and Identified Regulation, r_s (30)=0.52, p<0.001; External Regulation, r_s (30)=0.43, P<0.05; Sympathy toward Disabled People, r_s(30)=0.48, p<0.001; and SRA, r_s (30)=0.40, p<0.05, and a significant negative correlation with Amotivation r_s (30)= -0.38, p<0.05.

Table 6.8 Correlations between SIMS, IDP, SRA scales and the number of images described in the two conditions

Measures	InMot (N=37)		ExMot (N=30)	
	r_s	**(sig)**	**r_s**	**(sig)**
SIMS subscales				
Intrinsic Motivation	0.17	n.s.	0.08	n.s.
Identified Regulation	0.18	n.s.	0.52	0.003
External Regulation	0.18	n.s.	0.43	0.019
Amotivation	-0.01	n.s.	-0.38	0.037
IDP Scale				
Sympathy	-0.17	n.s.	0.48	0.007
Discomfort	0.00	n.s.	-0.27	n.s.
SRA Scale	-0.21	n.s.	0.40	0.028

In addition, to investigate whether there was a relationship between the number of images produced and students' ratings to their likelihood of participation in the DescribeIT project (measured in Phase 1) a Spearman correlation was conducted. It was only possible to match 31 students who had participated in both Phases 1 and 2.

(this was because the rest of students did not provide their emails in one of the data collection phases, which made it impossible to match their participation in Phase 1 to Phase 2). The results showed a significant positive correlation between the number of images described by this group of students and their ratings of likelihood of participations, r_s (31)=0.46 ,p<0.001.

6.4 Discussion and Conclusions

This study investigated how intrinsic and extrinsic motivational factors can affect the participation of students from two Arab countries in a crowdsourcing project to support blind and partially sighted students. The study explored both students' perception of what they think would motivate them to participate in the DescribeIT project, and students' behaviour in the DescribeIT project in relation to their situational motivation, attitudes towards people with disability, and their self-reported sense of altruism. In addition, the study investigated the quality of the image descriptions produced by students. For results summary see Tables 6.9, 6.10, and 6.11 end of this chapter.

In Phase 1 of the study, students showed a positive interest to participate in the DescribeIT project. Overall the students rated their likelihood of participating as positive (significantly above the midpoint of the scale), although they thought the likelihood of other students participating was only neutral. Students likelihood of participation ratings for themselves were significantly higher than their ratings of the participation of their peers. This is in line with the findings of studies 1, 2, 3, and 4 with students in the UK, which suggest the effect of the fundamental attribution error. As previously explained in Study 1 (See Chapter 4, section 4.4) the fundamental attribution error proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits.

When students were asked what they thought would motivate them and what would motivate others, they reported that wanting to help disabled students as the most important motivational factor for both themselves and other students. Their answers in the open-ended questions about this factor was also reflected in their ratings of the same motivational factor, in which altruism and wanting to help others factor was rated significantly above the neutral point. The importance of this factor was also reported by UK students both in the open-ended questions and their ratings in Studies

1, 2, 3, and 4. It also agrees with the findings of other research that investigated people's motivation to participate in crowdsourcing projects in general (e.g. Kuznetsov 2006, Oreg and Nov 2008) that found people's sense of altruism is one of the leading factors for self-reported motivations for participating in crowdsourcing projects such as Wikipedia.

In comparison to Studies 1, 2, 3, and 4 with UK students, a new motivational factor emerged from this study which is the idea of supporting research, especially research that could potentially improve the education of disabled students. Perhaps because Arab students are not as exposed as European students to the possibility of participating in research, students thought that they would be motivated by this factor, which could have influenced their actual participation to the DescribeIT project (in Phase 2 of the study). It is worth noting that in Phase 3 of Study 3 (See 5.4.4.3), when UK students were asked about why they did not participate in the DescribeIT project, being involved with other research was one of the reported reasons for not taking part in that study.

No-one explicitly mentioned the importance of the fun of the activity as a motivation to participate in DescribeIT project for themselves or other students. However, when students were asked to rate this factor, the results showed that students thought that the fun and entertainment of the activity is an important motivator, the result of the ratings is similar to the findings of Study 3 and other research (Brabham 2008, 2012, Hossain 2012, Kaufmann, et al. 2011) which reported the importance of this factor in self-report studies.

Improving one's skills was reported as significantly important by students of the present study, a finding which is in line with Brabham (2008, 2010, 2012) findings, in which he reported the importance of improving or developing artistic skills as a motivational factor. Students gave a few reasons for their ratings including that they use presentation slides all the time and participation in the DescribeIT project would improve their skills in knowing how to describe them effectively. It is interesting that some students thought that for them to be able to help disabled students they should improve their own skills in the first place. However, when students were asked what would motivate them and other students to participate to the DescribeIT project, only three students mentioned this factor. In addition, students throughout the other Studies 1,2, 3 and 4 with students in the UK were always neutral about this factor.

Students were neutral in their ratings about enhancing their job opportunities in the future and drawing attention to their skills. Students' thoughts on that were a mixture of wanting to improve their skills and hence, draw attention to these skills in their CVs, either to seek a job or to impress current employer. However, students who were less motivated by this factor thought gaining any personal outcome out of participation would not be their motivation, as they stressed again in their comments, they would be participating to help disabled students. This finding is contrary the findings of Brabham's (2008) study, in which he reported that some of iStockphoto members' motivation to participate to the platform is to improve marketability and to get a better job.

Interestingly, students reported the motivational factor of monetary reward significantly low, both when they were asked what would motivate them and when they were asked to rate this factor. Students thought it is inappropriate to gain money from helping disabled students. Whereas in Study 2, students in the UK were neutral about this factor, as some students explained they would consider participating for money only if they were in need of money. In the other hand, Studies 1, 3 and 4 with UK students, Lakhani et al. (2007) and Brabham (2008, 2010) all reported the importance of monetary reward as a self-reported motivation.

In addition, students thought that they would not be motivated by being in a competition with other students to participate to the DescribeIT project, this result is in line with the results of Studies 1, 2, and 4 with UK students. However, in Study 3 students were neutral about their feeling about this factor.

The motivational factors of social recognition and to pass the time were rated significantly low by students throughout Studies 1 to 5. This suggests that students from different departments and different cultures think that receiving social recognition from helping disabled students is inappropriate. In addition, students in Studies 1 to 5 who all were full-time students thought that they could pass their time doing more enjoyable activities than describing images in digital learning resources. This particular finding about the passing the time factor is similar to Brabham's (2008) study, in which he reported that iStockphoto members do not participate to the iStockphoto to pass the time when they are bored.

In Phase 2 of the study, the results showed that students in general were intrinsically motivated to participate in the DescribeIT project both in the InMot and ExMot conditions. Students in both the InMot and the ExMot condition were neutral in their feeling that the description task as chosen by themselves, as measured by the SIMS sub-scale Identified Regulation. Perhaps this is because students were invited to participate in a crowdsourcing project, and they were not actively seeking to participate in crowdsourcing, as the demographic data showed that only two students had participated in crowdsourcing before.

The results on the External Regulation sub-scale were very interesting. While it was understandable that students in the InMot condition were not motivated by external rewards or negative consequences, it was not expected that students in the ExMot condition would not be motivated by the external motivator of improving their skills. The effect of the external motivation factor seems to be substantially weakened in the ExMot condition because students were significantly intrinsically motivated to participate in this particular project to support disabled students.

The level of participation was measured by the number of images described by each student. It was hypothesized that there would be a positive relationship between the number of images described and the students' sense of altruism, attitudes towards disabled people, and their situational motivation. The results showed that there was no difference in the number of images produced between the two conditions. In terms of the number of images produced by each student, there was a significant relationship between three of the situational motivational factors and the number of images for the ExMot condition, with positive relationships between number of images and Identified Regulation, External Regulation and a negative relationship with Amotivation. The first and third correlations make sense. Students who felt they chose to participate voluntarily (as measured by the Identified Regulation sub-scale) described more images, and those who did not value the task by any means (as measured by the Amotivation sub-scale) described fewer images. However, the second correlation, that of a positive relationship between number of images described and External Regulation, which suggests an extrinsic motivation for doing the task, seems at odds with other findings in this study which suggest intrinsic motivations for doing the task. For example, there was a positive correlation between number of images description and the Sympathy sub-scale of the IDP. In addition, this group of students reported that they are positively altruistic, as measured by the SRA. Unexpectedly, there was no relationship between the number of images described and any of these variables in the InMot condition.

Lastly, as another way of measuring students' level of participation with the DescribeIT Project, the quality of the image descriptions produced was measured. Typically, a useful image description for blind and partially sighted people needs to be between 80 and 120 words long to have enough detail about the contents of the image (Chen, 2013). It was also found by the positive significant relationship between the number of object mentioned in a description and the word count of the descriptions, meaning the longer the description is, the more detail it has. The results of the present study showed that 74% of the descriptions produced were between 1 and 50 words long, which means that these descriptions are probably not detailed enough for blind and partially sighted students.

To conclude, Arab students showed an interest in participating in the DescribeIT project when asked about their likelihood of participations. These students thought that wanting to help others is the most important motivational factor to participate in the socially responsible crowdsourcing project, DescribeIT. In comparison to previous Studies (1 to 4) with UK students, a new motivational theme emerged, that of wanting to support research. This was reflected in students' participation in Phase 2 of the study. Phase 2 of the study showed that students were intrinsically motivated to describe images to blind and partially sighted students. The intrinsic motivation even dominated the effect of the extrinsic motivational factor in the extrinsic condition. A significantly positive relationship was found between students' ratings of their likelihood of participation (in Phase 1) and the number of images they produced (in Phase 2). The quality of the descriptions produced was not high as one would wish for, as most of the descriptions were not long enough to have enough detail of the contents of the images. However, these could be used as short descriptions.

The results of this study showed very interesting results in both Phase 1 and 2. The fact that Arab students participated in the DescribeIT project in far greater numbers in comparison to UK students, who in Studies 2, 3 and 4 had shown either a very low participation rate or a total lack of participation suggests that there are cross-cultural differences in motivations to support disabled students. Further research in cross-cultural differences in motivations to socially responsible crowdsourcing and the support of disabled students is needed. In addition, although the study was not primarily aimed to recruit women, the study ended up having far more women than men. Again, this could be investigated in future work to see if there are gender differences in motivations to participate in socially responsible crowdsourcing projects.

Table 6.9 Summary of the results of Studies 1 to 5

	Study 1	Study 2	Study 3	Study 4	Study 5
Number of participants (phase 1)	271 students	27 students	26 students	36 students	92 students
Number of participants (phase 2)	N/A	8 students	phase 3: 19 students	2 students	89 students
Number of descriptions	N/A	86 descriptions	17 descriptions (round 2)	11 descriptions	584 descriptions
Likelihood of participations	Results of One-sample Wilcoxon tests				
Self	+**	n.s	+**	+**	+**
Others	n.s	n.s	+*	n.s	n.s
Motivational Factors					
Altruism	+**	+**	+**	+**	+**
Money	+**	n.s	+*	+**	-**
Academic credits	+**	n.s	+*	n.s	n.s
Job opportunities	+**	n.s	+*	n.s	n.s
Contributing	+**	n.s	n.s	+*	+**
Improving skills	n.s	n.s	n.s	n.s	+*
Fun/entertainment	n.s	n.s	+*	n.s	+**
Being connected	-*	n.s	+*	n.s	n.s
Attention	-**	-*	n.s	-**	n.s
Competition	-**	-*	n.s	-**	-**
Social recognition	-**	-*	-**	-**	-*
Pass time	-**	-**	-*	-**	-**

Table 6.10 Results of One-sample Wilcoxon tests for the SIMS scale of Study 5

SIMS Scale	InMot Condition	ExMot Condition
Intrinsic Motivation	+**	+**
Identified Regulation	n.s	n.s
External Regulation	-**	-*
Amotivation	-**	-**

Table 6.11 Correlation between number of images and SIMS, IDP, and SRA scale for Study 5

SIMS Scale	InMot Condition	ExMot Condition
Intrinsic Motivation	n.s	n.s
Identified Regulation	n.s	+**
External Regulation	n.s	+*
Amotivation	n.s	-*
IDP Scale		
Sympathy	n.s	+**
Discomfort	n.s	n.s
SRA Scale	n.s	+*

Chapter 7

Monetarily Rewards as a Motivation to Participate in a Socially Responsible Crowdsourcing Project

7.1 Introduction

Unfortunately, people with disabilities experience high levels of unemployment (Papworthtrust.org.uk, 2016), so they are not in a position to pay for services that make information accessible for them. In addition, it does not seem right that they should be expected to pay to access information that is freely available to non-disabled people. Therefore, the primary aim of this programme of research was to understand how best to motivate students to participate in crowdsourcing projects to support their disabled peers without using financial incentives. However, due to the low and sometimes total lack of participation in the DescribeIT project, it was decided to investigate if paying students to do the image description task would motivate them to participate to the DescribeIT project. Initially university students were recruited because they understood the context and educational intent of the images used in the digital learning resources. However, Studies 2, 3 and 4 have shown that for students in the UK, asking them to do the image description task for non-financial rewards does not work, although it did work for students in Arab countries. Therefore, the first study (Study 6) in this chapter will investigate whether giving students small monetary rewards will actually motivate them to participate in the project.

However, if it is necessary to pay students to create image descriptions, it may be easier, faster and cheaper to use an established micro-payment crowdsourcing platforms such as Amazon Mechanical Turk (MTurk) for the task. This would potentially maintain a steady pool of crowd members whenever needed for image description. A very successful example of using Mturk to recruit people to participate in socially responsible crowdsourcing project is the VizWiz application (Bigham et al. 2010) in which nearly real-time answers to questions about visual information are provided for blind people through a special service called quikTurkit which ensures the availability of MTurk workers when questions asked.

Nonetheless, if crowd members who do not understand the educational context of the images they are asked to describe, this may compromise the point of the project, so it is vital to investigate the quality of the descriptions they produce. Would they be as accurate and as useful as those produced by students? Therefore, the second study presented in this chapter (Study 7) recruited crowd members from MTurk for the DescribeIT project, with the aim of investigating whether these crowd members would undertake the image description task for a typical MTurk recompense and whether they would produce appropriate image descriptions and of comparable quality to the students.

7.2 STUDY 6: The Effect of the Extrinsic Factor of Monetary Reward on Students' Motivation to Participate in the DescribeIT Project

7.2.1 Method

7.2.1.1 Design

The study investigated the effect of the extrinsic motivation of monetary reward on students' participation in a crowdsourcing project to describe images for their blind and partially sighted peers. Levels of participation could then be compared to the effects of the non-monetary motivational factors used in Studies 2, 3 and 4.

Students received 20 pence (£0.20) per description. The dependent variables were the number of students who participated in the study and the number of images described. The quality of the image descriptions produced was also investigated to measure students' level of participation in the image description task.

In line with the previous studies, other dependent variables were measured using an online questionnaire, including:

1. Students' overall level of altruism was measured using the Self-Report Altruism Scale (SRA) (Rushton et.al. 1981), which measures students' self-reports of the frequency with which they have engaged in altruistic behaviours.
2. Students' overall attitude towards disabled people was measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994), which assesses students' perception of their interaction with disabled people. Maclean and Gannon's (1995) sub-scales were used which measure "Discomfort" and "Sympathy" toward people with disabilities.

3. Students' motivation while doing the image description task was measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000), to assess why students were engaged in the image description activity.

7.2.1.2 Participants

106 students were invited to take part in the study, 26 (24.5%) participated in the study. This comprised 7 women and 19 men. Their age range was 18 to 23 years old, with a mean age of 18.6 (SD = 1.3) years old. All students were from the UK, except for two from India, one from Bulgaria and one from Sweden. However, of those only 14 (54% of the 26 students) completed the online questionnaire, comprising 7women and 7men. Their age range was 18 to 23years old, with a mean age of 19.8 (SD= 1.8) years old. Students were from the same students (from a range of disciplines at the University of York) who had been invited to participate in Studies 2 , 3, and 4 but who had not participated in the DescribeIT project in those studies. This means that they had participated in Phase 1 but not in Phase 2 of the previous studies. The total number of students invited to participate in this study were: 22 students who participated in Phase 1 of Study 2, 26 students from Phase 1 of Study 3, 36 students from Phase 1 of Study 4, and 22 students from another study that is not reported in this thesis. The Phase 1 data for those final 22 students were included in the results of Study 1.

All the students reported using social media such as Facebook, Twitter and Snapchat. However, only seven students reported having participated in crowdsourcing projects before, such as Wikipedia, and Zooinverse.

Students were offered £0.20 per image description, to be given as Amazon gift vouchers. This was estimated based on the minimum wage in the UK[11] (£7.5 per hour) and they could earn around £5 of describing the images which seemed to be a reasonable reward. In addition, as an incentive to participate in the online questionnaire, students who participated were entered into a prize draw for one of 10 Amazon vouchers worth £10 each.

7.2.1.3 Materials

The online questionnaire was the same as one used in Phase 2 in Studies 3 to 5, and comprised four sections: Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000), Interaction with Disabled Persons Scale (IDP (Gething, 1994), Self-

[11] https://www.gov.uk/national-minimum-wage-rates

Report Altruism Scale (SRA) (Rushton et.al. 1981), and Demographic questions. For more details about each scale and the demographic questions. See Chapter 2, section 2.8 for details about the scales.

7.2.1.4 DescribeIT project

The DescribeIT project used in this study was the same one presented in Chapter 3, section 3.2. To help students create good image descriptions, the project information page provided guidelines on how to describe images for blind and partially sighted people and an example description of a typical image.

In the DescribeIT project for this study students were asked to describe images of PowerPoint slides from Professor Helen Petrie's inaugural lecture, titled: "Navigating in the dark: Technology for disabled and elderly people". In total, the project had 27 slides. The project was closed once the desired number of descriptions was reached, which was 385 descriptions (12-15 descriptions per slide). The use of inaugural lecture slides was decided for two reasons:

(1) Students of this study were students from range of disciplines at the University of York, and from previous experience it was time consuming to contact lecturers in different departments to get slides to use in the DescribeIT Project. So, it was appropriate to use slides that were created for a general public lecture which did not require any specific knowledge of the topic.
(2) The same pack of slides will be used in Study 7 with MTurk workers. This will allow the quality of the descriptions produced by students to be compared to those produced by MTurk workers for exactly the same set of images.

7.2.1.5 Data analysis: Image description quality

To assess the quality of the image descriptions, the number of words in the image description was used as a simple metric for the quality of the description. Typically, the longer a description, the more detail it has and the more useful it will be to visually disabled students.

In addition, the quality of about 10% of the descriptions collected in this study were assessed using two more detailed methods: a method developed from signal detection theory (SDT) and expert evaluation. The 10% of the descriptions was chosen randomly from the overall collected descriptions, excluding trivial descriptions.

Signal detection theory (SDT) method

Signal detection theory (SDT) (Snodgrass, Levy-Berger and Haydon, 1985) may be a useful approach to explore for the assessment of quality of image descriptions. Classically, SDT is used when one is investigating whether participants can detect stimuli against background noise. There are four possible outcomes (see Table 7.1): a *hit* (when the stimulus is present and the participant correctly detects it), a *miss* (when the stimulus is present but the participant fails to detect it), a *false positive* (when the stimulus is not present but the participant thinks it is), and a *correct rejection* (when the stimulus is not present and the participant perceives that correctly).

Table 7.1 The four possible outcomes in classic signal detection theory

	Stimulus detected	**Stimulus not detected**
Stimulus present	Hit	Miss
Stimulus not present	False positive	Correct rejection

This set of outcomes can be applied to the image description problem in the following ways (see Table 7.2). To create a good quality description, the describer should include as many important details of the image as possible (*hits* in SDT terminology), not miss any important details (*misses*), and not include any incorrect or inappropriate information (*false positives*). The idea of a *correct rejection* from SDT does not apply particularly well in the image description task, as there are an infinite number of things not in a particular image which the describer does not include. In addition, the number of hits should also be as high as possible for the number of words used (we can call this *efficiency*, although this is outside the classic SDT approach).

Table 7.2 The four possible outcomes in the signal detection theory method in relation to image description quality

	Item mentioned	**Item not mentioned**
Item in the image	HIT	Miss
Item not in the image	Error	N/A

Figure 7.1 Example of an image used in the DescribeIT project

This classification is applied to the description of Figure 7.1 to illustrate the approach:

- If an item is in the image and is mentioned in a description, this constitutes a *hit*. If the description mentions *girl*, *glasses*, or *watch* these are all *hits*.
- If an item in the image is not mentioned in the description this constitutes a *miss*. If the description fails to mention *books*, *head in hands*, *grey jumper*, these are all *misses*.
- If an item is not in the image but is mentioned in the description this constitutes a *false positive*. If the description mentions *laptop*, *sunglasses* or *boy* these are all *false positives*.

To measure the quality of the descriptions, these concepts from SDT were used to derive the following formulae to assess different measures of image description quality:

Accuracy reflects the degree to which the image description is precise in relation to the content of the image. The term *items available* refers to all the items and attributes in the image that could be described. The set of guidelines provided to participants in the DescribeIT project was used as the baseline for calculating the *items available.* For example, what objects, people, locations are shown in the image, what are the interesting and important features of objects in the image, the colour of objects, their size, orientation, and relationships between them.

A high quality description has a high accuracy score and a poor quality description has a low accuracy score.

Accuracy = $\left[\frac{Hits}{Total\ of\ items\ avaliable}\right] \times 100$

Error score reflects the frequency false positives in an image description. A high quality description has a low error score and a poor quality description has a high error score.

Error Score = $\left[\frac{False\ positives}{Total\ of\ items\ mentioned}\right] \times 100$

Efficiency is the extent to which a description provides accurate information in the least number of words. The term *word count* refers to the number of words in the image description.

A high-quality description has a high efficiency rate and a poor-quality description has a low efficiency rate.

Efficiency = $\left[\frac{Hits}{Word\ count}\right] \times 100$

Expert evaluation

Eight experts were asked to assess whether or not the descriptions would be useful for blind and partially sighted students. All the experts were Native English speakers and had over 10 years of experience working in the accessibility field and with blind and partially sighted people. The experts were provided with a copy of the same set of guidelines given to participants on how to describe images for blind and partially sighted people. Each expert evaluated 10 images, with two descriptions each (20 descriptions in total). Each set of 20 descriptions was given to two experts. They were asked to:

1. Decide whether the description was appropriate for a short or long description
2. Rate the overall quality of the description on a 5 point Likert item (1= very poor to 5= excellent).

3. Provide a short explanation of their rating.
4. Rate their confidence in their decision on a 5 point Likert item (1= Not at all confident to 5= very confident).

The resolution system for the experts' assessment who had the same set of descriptions was as follows:

1. If the two experts agreed on whether the description is short or long, then no resolution is needed.
2. If the two experts disagree on whether the description is short or long then the decision was chosen based on the level of confidence of the two experts involved. The decision of the one who had given the higher confidence was chosen.
3. If the two experts disagree on whether the description is short or long and both had the same confidence level then the description was not included in the data analysis.

The quality of the description rating was the mean of the ratings from the two experts.

7.2.1.6 Pilot Study

A pilot study was conducted to check for any technical problems with the version of the DescribeIT project to be used in this study, with one PhD student from the Department of Electronic Engineering. No technical problems were detected. That student was excluded from the main study.

7.2.1.7 Procedure

A recruitment email was sent out to the 106 students who participated in Phase 1 of Studies 2, 3, and 4. One reminder email was sent several days later. The project was closed once the number of descriptions required has been reached.

7.2.1.8 Data Analysis

The data collected was not normally distributed for some of the variables. Thus, it was decided to use nonparametric tests for all analyses. It is worth noting that the same results were found when parametric tests were used with the normally distributed variables.

7.2.2 Results

A total of 26 students participated in the DescribeIT project, produced 385 descriptions, a mean of 17.5 (SD=9.1) descriptions per student. However, only 14 students completed the online questionnaire. This subset of students produced a total of 299 descriptions, a mean of 21.4 (SD= 8.4) descriptions per student. Interestingly, most of the students (18 students, 69%) participated within the first two hours of sending the recruitment email, the first few descriptions were received two minutes after sending the recruitment email.

The mean and median ratings and standard deviations of students' scores on the four sub-scales of the SIMS, the two sub-scales of the IDP, and their overall SRA scores are summarized in Table 7.3.

Table 7.3 Mean and median ratings and standard deviation of SIMS, IDP, SRA sub-scales, and the results of one-sample Wilcoxon signed rank tests

Measure	Mean (SD)	Median	W	*P*
SIMS Subscales				
Intrinsic Motivation	3.3 (1.1)	3.1	- 2.0	0.044
Identified Regulation	3.6 (1.1)	3.4	- 1.2	n.s
External Regulation	1.6 (0.8)	1.3	- 3.3	0.001
Amotivation	2.2 (1.4)	1.5	- 2.9	0.003
IDP Scale	66.2 (10.2)	71	1.4	n.s.
IDP Sub-scales				
Sympathy	21.7 (4.0)	23.0		
Discomfort	11.6 (4.0)	12.0		
SRA Scale	38.6 (9.7)	43.0	-1.8	n.s.

One-sample Wilcoxon signed rank tests were conducted to investigate whether the scores on each of the SMIS sub-scales were above the midpoint of 4 or not. As shown in Table 7.3, the scores for Intrinsic Motivation, External Regulation, and Amotivation were significantly lower than the midpoint of the scale. Whereas, scores for Identified Regulation were not significantly different from the midpoint.

The mean IDP score was 66.2 (SD= 10.2). A comparison with Brown et al. (2009) reported results on the IDP for British undergraduate students was made, in which their mean IDP score was 62.31. A one-sample Wilcoxon Signed Rank test found that the scores of the current group of students is not significantly different from those of Brown et al. (2009). On the Sympathy, sub-scale students scored a mean of 21.7 (SD=4.0) and on the Discomfort sub-scale students scored a mean of 11.6 (SD=4.0).

On the SRA students scored a mean of 38.6 (SD=9.7). A one-sample Wilcoxon Signed Rank test found that the scores of the current group of students is not significantly different from those of scores provided by Rushton et al. (1981)

Spearman correlations were calculated to investigate whether there was a relationship between students' scores on the four SIMS sub-scales and the number of images they described. There were no significant correlations. The results summarized in Table 7.4 A further two Spearman correlations were calculated between the scores on the two sub-scales of the IDP and the numbers of images described. There was no significant correlation with Sympathy, but a significant negative correlation with discomfort $r_s(14)=-0.5$, $p=0.05$. Finally, a Spearman correlation was calculated between participants' scores on the SRA and the numbers of images described. Again, there was no significant correlation.

The relationship between students' ratings to their likelihood of participate in the DescribeIT project which they had given in Phase 1 of the study (data collected in Studies 2, 3, and 4) and the number of images described was also investigated. The Spearman correlation test was conducted showed no significant correlation. Another Spearman correlation was calculated between students' ratings to the monetary reward motivational factor and the number of images described. There was no significant correlation. In addition, to investigate if students' ratings of the monetary reward factor were significantly above the midpoint of the rating scale, a one-Sample Wilcoxon Signed Rank test was carried out. This showed that students' ratings of monetary reward factor were significantly higher than the neutral midpoint ($W= 3.0$, $p<0.01$).

To investigate the performance of the students, the quality of the descriptions was examined. The mean number of words per description was 44.0 (SD= 33.8) and the word count ranged from 2 to 229 words. The range of the word counts is illustrated in Table 7.5

Table 7.4 Correlations between SIMS, IDP, SRA scales and the number of images described

Measure	Mean (SD)	Median	r_s	*P*
SIMS Subscales				
Intrinsic Motivation	3.3 (1.1)	3.1	-0.21	n.s.
Identified Regulation	3.6 (1.1)	3.4	0.06	n.s.
External Regulation	1.6 (0.8)	1.3	-0.36	n.s.
Amotivation	2.2 (1.4)	1.5	-0.41	n.s.
IDP Scale				
Sympathy	21.7 (4.0)	23.0	-0.02	n.s.
Discomfort	11.6 (4.0)	12.0	-0.5	0.050
SRA Scale				
Self-Report Altruism	38.6 (9.7)	43.0	-0.04	n.s.

Table 7.5 Frequency and Percentage of the Word Count Ranges in the Image Descriptions.

Word Count Range	Frequency	Percentage
1 - 50	266	69%
51 -100	92	24%
100+	27	7%
Total	**385**	**100%**

Further detailed quality assessment was carried out with 40 image descriptions, using the SDT formulae and experts' evaluation. The scores of SDT formulae for each of accuracy, error score, and efficiency are summarized in table 7.6. The accuracy of the images descriptions showed a mean of 49.3% (SD=31.0) which represents the degree to which the image description is precise in relation to the content of the image. The error rate or the frequency of incorrect items mentioned in an image description showed a mean of 0.0% (SD=0.0). Lastly, the efficiency of the image descriptions, which referred to their ability to provide a quality description with the least number of words showed a mean of 21.0% (SD=11.0).

Table 7.6 Mean, median and standard deviation of accuracy, error score, and efficiency of the image descriptions

SDT Measures	Mean	Median	SD
Accuracy	49.3%	44.0%	31.0
Error score	0.0%	0.0%	0.0
Efficiency	21.0%	21.0%	11.0

In the expert assessment, experts rated the quality of the descriptions on a 5-point Likert item (1=Very poor, 5=Excellent), this showed a mean of 3.2 (SD=1.0). A one-sample Wilcoxon signed rank test was conducted to investigate whether the experts' ratings were above the midpoint of 3 or not. The results showed that their ratings were not significantly different from the neutral midpoint of the scale. Table 7.7 represents some examples of experts' comments about the quality of the image descriptions with the associated rating.

Table 7.7 Examples of experts' comments about their decision and quality ratings for the image descriptions

Rating	Experts' comments
1	*The text in the image is not described and the key features of the image are not included.* (Expert 7)
2	*Describes entire slide and several pictures. Not terribly bad in detail* (Expert 3)
3	*This description is good but some important details are missing, such as the number of magazine covers and their arrangement* (Expert 6)
4	*A very good description that includes all of the necessary details. It is a little vague ("appears to have come"), does not mention that it is a black and white image, and does not need "The image here is of" but is a reasonable attempt.* (Expert 4)
5	*long text is comprehensive and gives positioning of the two photos, and details of clothing and activities of the people depicted.* (Expert 1)

A Spearman correlation was conducted to determine if there were any relationships between accuracy level and experts' ratings. The results showed that there was a significant positive correlation between the accuracy and experts' ratings, $r_s(40)=0.5$, $p<0.01$.

A Mann-Whitney U test was carried out to compare experts' decision about whether a description was more appropriate for a short or long description with the efficiency level of the description. The results showed that the efficiency level in the short descriptions were significantly higher than in the long descriptions ($U = 59$, $p<0.05$).

7.2.3 Discussion and Conclusions

This study investigated the effect of the extrinsic motivational factor of small monetary rewards on students' participation in a crowdsourcing project to describe images for their blind and partially sighted peers. In addition, it explored students' behaviour, namely the number of images they described, in the DescribeIT project in relation to their situational motivation, attitudes toward people with disability, and their self-reported sense of altruism. The performance of students, the quality of the image descriptions produced, was also investigated as another measure of students' participation in the descriptions task. For results summary see Tables 7.12, 7.13 and 7.14 end of this chapter.

The scores on the Intrinsic Motivation sub-scale of the SIMS show that students were not intrinsically motivated to participate in the DescribeIT project. This is understandable as this group of students was invited to participate in the project (using intrinsic or non-financial extrinsic factors) previously in Studies 2, 3, and 4 but they did not participate, despite reminder emails sent to them. However, when they were offered the extrinsic motivation of money, not only did they participate, but they also very often participated within the first few hours of sending the first recruitment email. In addition, students were neutral in their sense that they had voluntarily chosen to participate in the DescribeIT project (the Identified Regulation sub-scale of the SIMS). This finding agrees with the findings of Study 5, in which Arab students in both the intrinsic and extrinsic motivation conditions were neutral in their sense that one has voluntarily participated (as measured by Identified Regulation). Perhaps this is because students did not intentionally seek to participate in a socially responsible crowdsourcing project to support disabled students, rather they were invited to take part in the project.

However, the finding on the external regulation subscale also interesting. The results showed that students' ratings were significantly below the neutral midpoint of this scale. Although this result agrees with the results of study 5, this result was more surprising with this particular group of students than with those of the extrinsic motivation condition in Study 5. Firstly, the students in the extrinsic motivation condition of study 5 were significantly intrinsically motivated to participate to the DescribeIT project, which could have resulted in weakening the extrinsic motivational factor. Whereas, students in this study were not intrinsically motivated. Secondly, considering the fact that students in this study only participated when they were offered money does not match their results on the External Regulation subscale. This finding suggests that social desirability bias might have influenced students' answers. Students might have thought it against social norms to be paid to help disabled peers, so they found it difficult to admit or believe that their motivation to help disabled students is mainly because of the monetary reward. However, it is worth noting that from the total 26 students who participated in this study only 14 students (54%) participated in the online questionnaire, which is relatively small number of participants.

Another unexpected result, was on the Amotivation subscale. The results of the SIMS scale showed that students were neither intrinsically nor extrinsically motivated to participate in the DescribeIT project. This means that they did not value the purpose of the project nor the reward expected as outcome of their participation. Yet, the results suggest that students were significantly not Amotivated in their behaviours.

Students' level of participation was firstly measured by the number of images described by each student. A positive relationship between the number of images described and the students' sense of altruism, attitudes towards disabled people (discomfort and sympathy), and their situation motivation was hypothesized. However, only a negative relationship was found between the number of images and discomfort subscale. The more discomfort students felt towards disabled people, the fewer images they described.

In addition, as another way of measuring students' level of participation in the description task, the quality of the image descriptions produced was measured. Typically, a useful image description for blind and partially sighted people needs to be between 80 and 120 words long to have enough detail of the contents of the image. The results of the present study showed that approximately 69% of the descriptions produced were between 1 and 50 words long, which means that these descriptions are probably not detailed enough for blind and partially sighted students. However, some of these descriptions could be suitable to be short descriptions.

The further analysis of the quality of the descriptions showed a moderate accuracy level. This can be explained by experts' comments, in which they noted that students had missed key details of the images in their descriptions. There was a significant correlation between accuracy and experts' ratings, this shows that the more information provided the more the experts thought the descriptions useful. However, the descriptions produced by students had no false positive information (error score of 0.0%). This means that students missed mentioning some key details of the images, but at least they did not mention things which were not presented in the image. The efficiency level of the descriptions provided was very low, meaning that students could need help in learning how to create more efficient descriptions. This is very important to blind and partially sighted students, as they are probably listening to the descriptions with a screen reader, which is inherently slower than visual reading. So making a text as efficient as possible is important. There was a higher level of efficiency in the descriptions chosen by experts as short descriptions, which is an encouraging sign.

In conclusion, students were motivated to participate in the DescribeIT project by the extrinsic motivational factor of small monetary rewards., although the results of the subscale of the SIMS scale external regulation showed quite the opposite. This is another indication about the importance of not relying solely on self-reports of motivation when assessing motivation. However, although paying students motivated them to participate the quality of their descriptions was average.

7.3 STUDY 7: Participation in the DescribeIT Project by Amazon Mechanical Turk Workers

7.3.1 Introduction

Amazon Mechanical Turk (MTurk) is an open online marketplace that allows crowdsourcers to outsource tasks to crowd members who are paid upon successful completion of each task. Although MTurk is a relatively new crowdsourcing platform it was chosen for this study as it has been used extensively in many research studies (e.g. Chandler and Kapelner 2010; Ross et al. 2010; Ipeirotis 2010; Rogstadius et al. 2011; Thaler, Simperl, and Wölger 2012; Mao, Kamar and Horvitz 2013; Goncalves et al. 2013; Difallah et al. 2015) In addition, the MTurk platform is the most well studied crowdsourcing platform. The key strength of any successful crowdsourcing platform such as MTurk is the ability to access a large pool of people who are willing to do tasks for relatively little monetary reward. This has led many researchers to recruit participants through MTurk to participate in online research, including behavioural research. Many studies cited by Mason and Suri (2012) have shown that the behaviour of workers in MTurk is comparable to the behaviour of similar individuals offline or in other online contexts.

7.3.1.1 Background: The Motivations of MTurk Workers

This section will focus on research that has been conducted to investigate what motivates people to participate in Amazon Mechanical Turk (MTurk), as it is the platform used to deploy DescribeIT in the present study.

Ipeirotis (2010) assessed the motivation of workers on Amazon MTurk. The workers were mostly from the USA and India, approximately 50% and 40% respectively. Ipeirotis conducted an online questionnaire with 1,000 workers asking them to respond to six statements about why they completed tasks in MTurk. Workers were also given the chance to elaborate more on why they participate on MTurk, by answering open-ended questions. 27% of the Indian workers reported that they were motivated by the monetary rewards on MTurk, as these were their primary source of income, whereas only 12% of American workers reported that MTurk was their primary source of income. However, only 37% of the Indian workers reported participating on MTurk to earn additional money (i.e. as a secondary source of income) compared to approximately 60% of the American workers. Perhaps the most interesting finding was that about 70% of Americans and roughly about 60% of Indian workers agreed that "Mechanical Turk is a fruitful way to spend free time and get some cash". Nonetheless, Paolacci, Chandler, and Ipeirotis (2010) noted that not all workers participate for money, as many of them participate for intrinsic motivational factors such as for fun and entertainment (reported by 41% of workers), and to pass the time (reported by 32%).

Ross et al. (2010) also investigated what motivates the workers on Amazon MTurk with 733 workers, by using an online questionnaire. They found that 5% of the American workers relied on MTurk as a source of income to "make basic ends meet" compared to 13% of the Indian workers. Whereas 50% of the American workers thought that money earned through MTurk is "nice but does not change their circumstances", in comparison to 31% of the Indian workers. Only 24% of the

American workers agreed when they were asked if the money is a way to pay for extras, whereas 32% of the Indian workers agreed. However, when the workers were asked how important the money earned on MTurk was to them, only 12% of American and 10% of Indian indicated that “money is irrelevant”.

The findings of the previous studies (Ipeirotis 2010; Ross et al. 2010) agree with the findings of Horton, Rand, and Zeckhauser (2011) who conducted an online study with 302 workers and reported that money was the primary motivation to participate in MTurk in general. However, they noted that workers also valued other motivational factors, for example Indian workers reported that they “want to learn new skills”, while Americans reported that they “want to have fun”.

It is interesting that in most of the studies that investigated MTurk workers' motivations, money was always reported as the primary motivational factor, in spite of the very low payment rate (under $2 /hour); workers in general often earn less than $20 per week (Ipeirotis, 2010a). Martin et al. (2014) highlighted the ongoing debate and extensive discussions regarding whether MTurk workers are really motivated by money or by other factors. He argued that perhaps what makes it difficult for researchers to believe that workers are participating for the monetary reward is the low payment for completing tasks.

The results of Study 6 have suggested that students were motivated to do the image description task for monetary rewards. However, if paying students is necessary to do the image description task, it may be faster and cheaper to use Amazon Mechanical Turk (MTurk) for the task. With this in mind, Study 7 aimed to investigate whether MTurk workers would undertake the image description task for a typical MTurk recompense and whether they would produce image descriptions of appropriate quality and comparable quality to the students. In addition, the study investigated the relationship between MTurk workers' participation and performance in relation to their situational motivation, their attitude towards people with disabilities, and lastly to their self-reported sense of altruism.

7.3.2 Method

7.3.2.1 Design

The study explored the behaviour of MTurk workers participating in the DescribeIT project to describe images for blind and partially sighted students. The main dependent variables were the number of participants, the number of images described per participant and the quality of the images descriptions produced.

The other variables were:

As with the previous studies, participants' overall self-reported level of altruism, measured using the Self-Report Altruism Scale (SRA) (Rushton et al. 1981); participants' attitudes towards people with disabilities, measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994); participants' motivation while doing the image description task, measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000). This set of variables plus demographic information were collected through an online questionnaire.

7.3.2.2 Participants

164 MTurk workers participated in the DescribeIT project, describing images. However, only 86 completed the online questionnaire and of those 10 participants did not answer the demographic questions. Thus, we have demographic information about 76 workers. This comprised 44 women and 32 men, aged 19 to 67 years, with mean age of 38.3 years (SD = 10.8). Participation was restricted to a minimum level of MTurk "qualification", that is to workers who have a Human Intelligence Task (HITs) approval rate of greater than 95% and who have greater than or equal to 5000 approved HITs (this approval rate is used to restrict participation to workers who have demonstrated good skill and reputation to carry out tasks).

Workers were offered USD 0.10 (equivalent GBP 0.08) as payment per image description; this was in line with other MTurk research tasks of the same magnitude. In addition, as an incentive to complete the online questionnaire, participants were offered a USD 0.40 (GBP 0.32) bonus.

7.3.2.3 DescribeIT Project

A version of the DescribeIT project was developed on the Amazon MTurk platform. To help sighted people create good image descriptions, the DescribeIT project information page provided guidelines [Appendix A.1] on how to describe images for blind and partially sighted people and an example description of a typical image. The design of the project was similar to those used in Studies 2 to 6. Once workers accepted the HIT they were presented with a slide and a text-box in which to type their description of the image on the slide.

The same set of images used in Study 6, PowerPoint slides from Professor Helen Petrie's inaugural lecture, titled: "Navigating in the dark: Technology for disabled and elderly people" were used.

7.3.2.4 Materials

The online questionnaire was the same one used in Phase 2 questionnaire in Studies 3 to 6, and comprised of four sections, including: Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000), Interaction with Disabled Persons Scale (IDP (Gething, 1994), Self-Report Altruism Scale (SRA) (Rushton et.al. 1981), and Demographic questions. For more details about each scale see Chapter 2, section 2.8.

7.3.2.5 Data analysis: Image description quality

The same measures of image description quality were used as in Study 6: word count, SDT method and expert assessment. For the expert assessment, each expert evaluated 10 images. There were two descriptions of each image, thus each expert assessed 20 descriptions in total (80 descriptions in total). The experts were the same ones who participated in Study 6.

7.3.2.6 Procedure

The project was created on Amazon MTurk, and was available for 5 days. A link to the online questionnaire hosted on QuestionPro[12] was posted on the project information page for workers to follow. Workers were asked to describe the images with the help of the guidelines.

7.3.2.7 Data Analysis

Most variables in the data collected were not normally distributed, except for SRA scale scores, so it was decided to use nonparametric tests for all analyses.

7.3.3 Results

164 workers produced 851 descriptions, a mean of 5.6 (SD=8.5) descriptions per worker. However, the 75 workers who completed the online questionnaire, produced a total of 501 descriptions, a mean of 6.7 (SD=9.3) descriptions per worker.

[12] https://www.questionpro.com/

The mean and median ratings and standard deviations of workers' scores on the four sub-scales of the SIMS, the two sub-scales of the IDP, and their overall SRA scores are summarized in Table 7.8

One-sample Wilcoxon signed rank tests were conducted to investigate whether scores on each of the SIMS sub-scales were above the midpoint of 4 or not. The results are summarized in Table 7.8. Scores for the Intrinsic Motivation and Identified Regulation sub-scales were significantly higher than the midpoint of the scale ($p<0.00$ in both cases). On the other hand, scores on the Extrinsic Regulation and Amotivation sub-scales were significantly lower than the midpoint of the scale ($p<0.00$ in both cases).

Spearman correlations were calculated to investigate whether there was a relationship between workers scores on the four SIMS sub-scales and the number of images they described. Results are summarized in Table 7.8. There were no significant correlations. A further two Spearman tests were calculated between the scores on the two sub-scales of the IDP and the numbers of images described; these also showed no significant correlations. Finally, a Spearman's correlation test was calculated between workers' scores on the SRA. Again, there was no significant correlation and the numbers of images described.

Table 7.8 Mean and median ratings and standard deviation of SIMS, IDP, SRA sub-scales, the results of one-sample Wilcoxon signed rank tests and Spearman correlations of SIMS, IDP, SRA sub-scales and number of images.

Measure	Mean (SD)	Median	W	*p*	r_s	*p*
SIMS Subscales						
Intrinsic Motivation	5.3 (1.1)	5.5	6.9	0.000	0.7	n.s
Identified Regulation	5.1 (1.0)	5.0	6.6	0.000	0.3	n.s
External Regulation	2.8 (1.8)	2.3	-5.4	0.000	0.2	n.s
Amotivation	2.3 (1.5)	1.9	-6.9	0.000	0.7	n.s
IDP Scale	72.6 (13.1)	73.0	5.4	0.000		
IDP Sub-scale						
Sympathy	24.1 (3.3)	24.0			0.2	n.s
Discomfort	13.0 (5.8)	12.0			0.7	n.s
SRA Scale	51.1 (10.2)	52	5.1	0.000	0.1	n.s

The mean IDP score was 72.6 (SD= 13.1). A one-sample Wilcoxon Signed Rank test found that the scores of MTurk workers was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=5.4$, $p<0.001$). On the Sympathy, sub-scale students scored a mean of 24.1 (SD=3.3) and on the Discomfort sub-scale students scored a mean of 13.0 (SD=5.8).

On the SRA, MTurk workers scored a mean of 51.1 (SD=10.2). A one-sample Wilcoxon Signed Rank test found that the scores of MTurk workers was significantly higher than the median of SRA scores provided by Rushton et al. (1981) ($W=5.1$, $p<0.001$).

To further investigate the performance of the MTurk workers, the quality of the descriptions was investigated as a further measure of workers' participation in the description task. The mean number of words per description was 37.8 (SD=40.5) and the word count ranged from 1 to 372 words. The range of the word counts is illustrated in Table 7.9.

Table 7.9 Frequency and Percentage of the Word Count Ranges in the Image Descriptions.

Word Count Range	**Frequency**	**Percent**
1 - 50	652	77%
51 -100	138	16%
100+	61	7%
Total	**851**	**100%**

The mean and median ratings and standard deviations of the accuracy, error rate, and efficiency of the image descriptions are summarized in Table 7.10. Overall, the accuracy of the images descriptions showed a mean of 52.9% (SD = 29.2) which is the degree to which the image description is precise in relation to the content of the image. The error rate or the frequency of incorrect items mentioned in an image description showed a mean of 1.6% (SD=4.7). The results of both accuracy and error rate suggest that workers generally succeeded in mentioning in their descriptions items related to the content of the images, but failed to mention some important details, and the frequency of mentioning incorrect items in their descriptions was very low. Lastly, the efficiency of the image descriptions, which refers to their ability to provide a quality description with the least number of words showed a mean of 14.3% (SD=15.5), this suggest that the descriptions were unnecessarily long.

In the expert assessment, experts' ratings to the quality of the descriptions showed a mean of 3.2 (SD=1.1) on a 5-point Likert item (1=Very poor to 5=Excellent). The mean and median rating and standard deviation of expert assessments are summarized in Table 7.11. On a 5-point Likert item (1=Not at all confident to 5=Very confident) experts rated their confidence in their decisions with a mean of 3.9 (SD=0.7).

Table 7.10 Mean, median and standard deviation of accuracy, error score, and efficiency of the image descriptions

SDT Measures	Mean	Median	SD
Accuracy	52.9%	43.8%	29.2
Error score	1.6%	0.0%	4.7
Efficiency	14.3%	9.5%	15.5

Table 7.11 Mean, median and standard deviation of experts' ratings on the quality of the image descriptions and their confidents on their decisions

Experts Measures	Mean	Median	SD
Experts' confident	3.9	4.0	0.7
Experts' ratings	3.2	3.5	1.1

Experts comments about why they gave their specific ratings gave more insight to explain the results obtained by the SDT method. For example, the result of the accuracy of the image descriptions (Mean=52.9%, SD=29.2) suggest that the workers have mentioned some important details, but also missed important aspects. In their evaluation experts noted that, when for example some of them said:

It gives quite a lot of good detail, but misses out on one or two things that are important (Expert 1)

Description mentions the type of graph and parameters but omits important data such as the period of time and the opposite trends (Expert 7)

Similarly, the efficiency level of the descriptions (Mean= 14.3%, SD=15.5) was low, mainly because workers mentioned unnecessary information in their descriptions. The following are some experts' comments to illustrate this point:

*It contains good information, but too much waffle as well (*Expert 1)

*Very detailed description, but the wealth of detail overwhelms the reader. Cutting it down a lot and leaving the key details about the group and location would make this more useful (*Expert 3)

The description covers the text on the image in some detail which seems unnecessary, and that more attention should have gone into describing the image itself. (Expert 8)

A series of Spearman correlations were conducted to determine if there were any relationships between accuracy level and experts' ratings. The results showed that there was a significant positive correlation between accuracy and experts' ratings (r_s (80)=0.47, p <0.01). A Mann-Whitney test was used to compare the experts' decision on whether a description was suitable for a short or long description with the efficiency scores of the descriptions. It showed that the efficiency scores of the descriptions which the experts decided were appropriate as long descriptions were significantly higher than for those descriptions which the experts decided were appropriate as short descriptions (U=198.5, p<0.001)

It is interesting that one worker who contacted the researcher personally to explain his main motivation of participating in the project. The worker said that the loss of his son due to a brain tumour which lead to his son's loss of the ability to see (among other things) was the key motivation, he said: *When he lost his vision I would talk with him constantly and describe what was going on around him to help give him some sense of normality. I understand completely the magnitude of the sense of sight and what it feels like to loose it from his perspective.* (J). Unsurprisingly, that worker completed all the available HITs with good description quality and completed the online questionnaire.

7.3.4 Discussion and Conclusions

This study of Amazon Mechanical Turk workers participation in a crowdsourcing project to describe images for blind and partially sighted students has investigated the relationship between MTurk workers' participation and performance in relation to their situational motivation, their attitudes towards people with disabilities, and lastly to their self-reported sense of altruism. The quality of the descriptions produced was also investigated. The study has yielded some interesting but unexpected results. For results summary see Tables 7.12, 7.13 and 7.14 end of this chapter.

The study showed that MTurk workers were intrinsically motivated to participate in the DescribeIT project, which was similar to the results of Study 5. However, in contrast to the findings of Studies 5 and 6, MTurk workers perceived the description task as chosen by themselves, in the sense that one has voluntarily chosen to participate as measured by Identified Regulation sub-scale. Interestingly the results showed that workers were not externally motivated to participate in this particular project in spite of the monetary reward which was offered to complete the description task. This finding is in line with the findings of Studies 5 and 6.

The results of the present study on external regulation contrast to previous studies which used self-reports of motivation that report that money is the primary motivational factor to participate in MTurk (e.g Horton, Rand, and Zeckhauser 2011). This shows the importance of measuring both actual behaviour as well as self-reports of motivations. It suggests, as mentioned previously throughout the thesis, that self-report studies in the crowdsourcing area may be subject to social desirability effects. However, it is important to note that previous research have investigated the motivations of MTurk workers in general (i.e. not in a specific project). Whereas, this study investigated the motivation of MTurk workers in a specific project which would help disabled students to access online teaching materials. This is a characteristic which I believed would have increased workers' intrinsic motivation level to participate in this particular project.

Taking in consideration that the workers recruited in this study were deliberately chosen to be active members with high profile records in a popular platform like Amazon Mechanical Turk, it is understandable that the results showed a very low Amotivation scores, as the workers are highly motivated to participate in crowdsourcing in general. In addition, the study showed that workers self-reported general sense of altruism was higher than those reported by Rushton et al. (1981).

The level of participation was firstly measured by the number of images described by each worker, and it was hypothesized that there would be a positive relationship between the number of images described and the workers' sense of altruism, attitudes towards disabled people (both Discomfort and Sympathy), and their situational motivation. However, unexpectedly there was no relationship between participation in the DescribeIT project and any of these variables. It may have been that because this self-selecting sample was already high on intrinsic motivation, identified regulation and Sympathy for disabled people, there was not enough variation in these scores to show significant correlations.

The quality of the image descriptions produced was the second way of measuring workers' level of participation in the description task. For an image description to be useful for blind and partially sighted people it needs to be typically between 80 and 120 words long to have enough detail of the contents of the image. The results of the present study showed that approximately 77% of the descriptions produced were between 1 and 50 words long, which mean that these descriptions are probably not detailed enough for blind and partially sighted people. However, some of these descriptions could be suitable to be short descriptions.

Similar to the results of Study 6, the further analysis of the quality of the descriptions showed a moderate accuracy level. This can be explained by experts' comments, in which they noted that workers had missed key details of the images in their descriptions. The accuracy level correlated with experts' neutral ratings. However, the descriptions produced by workers had a slightly higher mean error score than the students in Study 6, as some workers had mentioned some objects which were not present in the images. The efficiency level of the descriptions provided was very low, this was mainly because workers mentioned unnecessary details in their descriptions. However, in contrast to Study 6 long descriptions were more efficient than short descriptions.

The fact that the worker (J) was motivated to participate because of his personal family circumstances suggest that family members or friends of people who are blind or partially sighted might have their own set of motivations that would drive them to participate in a socially responsible crowdsourcing projects.

It may be that describers such as MTurk workers need more instruction to gain more understanding of what is required in a description of an image for a blind or partially sighted person, particularly if the image is being used in education. Perhaps training would improve describers' skills to provide useful descriptions, which will be investigated in Study 10. The task of image description is a difficult one, and it may be that it needs to be broken down into smaller components, for example asking each person to only tag images instead of describing the whole images, which will be investigated in Studies 8 and 9.

Table 7.12 Summary of the results of Studies 2 to 7

Measures	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
Number of participants (phase 1)	27 students	26 students	36 students	92 students	26 students	Literature Review
Number of participants (phase 2)	8 students	No students	2 students	89 students	26 students	164 workers
Number of participants (phase 2 online questionnaire)	No students	phase 3: 19 students	1 student	67 students	14 students	86 workers
Number of descriptions	86 descriptions	17 descriptions	11 descriptions	584 descriptions	385 descriptions	851 descriptions

Table 7.13 Results of One-sample Wilcoxon tests for the SIMS scale of Studies 5 to 7

SIMS Scale	Study 5 (InMot)	Study 5 (ExMot)	Study 6	Study 7
Intrinsic Motivation	+**	+**	-*	+**
Identified Regulation	n.s	n.s	n.s	+**
External Regulation	-**	-*	-**	-**
Amotivation	-**	-**	-**	-**

Table 7.14 Correlation between number of images and SIMS, IDP, and SRA scale of Studies 5 to 7

SIMS Scale	Study 5 (InMot)	Study 5 (ExMot)	Study 6	Study 7
Intrinsic Motivation	n.s	n.s	n.s	n.s
Identified Regulation	n.s	+**	n.s	n.s
External Regulation	n.s	+*	n.s	n.s
Amotivation	n.s	-*	n.s	n.s
IDP Scale				
Sympathy	n.s	+**	n.s	n.s
Discomfort	n.s	n.s	-*	n.s
SRA Scale	n.s	+*	n.s	n.s

Chapter 8

Exploring the Effect of Task Type on Motivation to participate in a Socially Responsible Crowdsourcing Project

8.1 Introduction

This Chapter will present two studies which investigate how people's motivation and behaviour differed when they are asked to participate in a crowdsourcing project to *tag* objects in images in comparison to the task used in the previous studies in this programme of research, which was to *describe* images for blind and partially sighted students. The purpose of these studies was to firstly to investigate whether the task that had been used in the previous studies in this programme of research, that of describing images, was actually too hard and off-putting for students, and this contributed to the low rate of participation of the UK students. By making the task easier, in asking students only to tag objects in images, it is predicted that the participation rate will be higher. Therefore Study 8 asked students to participate in a crowdsourcing project, TagIt, very similar to the DescribeIt project used in the previous studies, but to simply tag objects in images. This allowed an investigation of the effect of the difference in task difficulty on participation by students.

Another explanation for the low rate of participation by British students is that the students invited to participate in the previous studies were not familiar with the concept of crowdsourcing. As reported in previous Chapters, few of the participants in the previous studies had participated in crowdsourcing projects (participation rates: Study 2, see section 5.3.2.2: 14%; Study 3: see section 5.4.2.2: 35%; Study 4: see section 5.5.2.2: 6%; Study 5, see section 6.2.2: 2%; and Study 6, see section 7.2.1.2: 27%). They may have not understood exactly how crowdsourcing works and have been hesitant about getting involved. Therefore, the second study in this chapter (Study 9) investigated the participation rate of crowd members on a crowdsourcing platform (Crowdcrafting.org) in both DescribeIT and TagIT projects. This allowed me to compare the participation rates of those not already involved with crowdsourcing with participants already involved in crowdsourcing and familiar with the concept, on both the original task of describing images and the simpler task of tagging objects in images.

The hypothesis is that the participation rates of the established crowd participators will be higher than the students in both cases.

The image tagging task is one of the most common tasks in crowdsourcing (Feyisetan et al., 2015). As discussed in Chapter 2, (see section 2.3), one of the successful examples of this task was the ESP game from Van Ahn and Dabbish (2004). They developed a system whereby sighted people could provide descriptive tags of images on the web which would be very useful for blind and partially sighted web users. Another example of the tagging task is the Peekaboom web-based game (von Ahn, Liu, and Blum, 2006) which used to locate objects in images.

The previous examples are based on making the task more enjoyable by implementing gamification aspects. One the other hand, LabelMe (Russell et al., 2008) is web-based tool for image tagging, in which the incentive used is to have access to the LabelMe database once you have tagged a certain number of images.

8.2 Study 8: Participation in the TagIT Project by Students

8.2.1 Introduction

To investigate whether the difficulty of the description task was responsible for the low participation of students in previous studies with UK students, in this study the image description task was simplified and students were asked to simply tag individual objects in images instead of providing whole descriptions. The tagging task could be genuinely useful in developing descriptions of images for blind and partially sighted students. Lecturers or students could use the tags to quickly and efficiently build up detailed descriptions. The aim of this study was to investigate the effect of task difficulty on students' motivation and behaviour, to establish whether asking students to do a simpler task increase their participation rate.

8.2.2 Method

8.2.2.1 Design

This study investigated participation in a crowdsourcing project to tag individual objects within images in comparison to the number of students who participated in the DescribeIT project which asked them to provide full descriptions of images (Studies 2, 3, and 4). The extrinsic motivation of improving students' own skills was included in the project information page and recruitment email. This particular motivational factor was chosen as it was investigated previously with the DescribeIT project in Study 2

(ExMot condition). The dependent variables were the number of students who participated in the project and the number of images tagged by each student. In addition, the quality of the image tags was measured. Similar to the previous studies, other variables were measured using an online questionnaire, including:

1. Students motivation while doing the tasks: was measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000).

2. Students overall attitude toward disability: was measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994).

3. Students overall level of altruism: which was measured using the Self-Report Altruism Scale (SRA) (Rushton et.al. 1981).

More information about the rationale for measuring these variables can be found in Chapter 2, section 2.8.

8.2.2.2 Participants

A total of 101 undergraduate and postgraduate students were invited to take part in this study. This included 30 undergraduate students in the Department of Theatre, Film and Television, and 71 MSc students in the Department of Computer Science, both at the University of York. Overall, 24 (24%) students participated in the study, these were all MSc. students. Of these, only 19 participated in the online questionnaire. 14 students completed the main questionnaire sections, and only 11 completed the demographic section. These comprised 5 women and 6 men. Their age range was from 21 to 35 years old, with a mean age of 27 (SD = 4.6) years old. Five students were from United Kingdom, three from India, and two from China (one preferred not to give their nationality).

8.2.2.3 Materials

A recruitment email and the project information page stated that: *Participating in this project will also improve your skills in analysing and tagging images which will make your blog articles, YouTube videos and images easy to find by search engines. Following our image tagging tips will help you learn how to provide good image tags.*

Similar to the previous studies, the online questionnaire consisted of four sections, including: *Self-Report Altruism Scale (SRA)* (Rushton et.al. 1981), *Situational Motivation Scale (SIMS)* (Guay, Vallerand, and Blanchard, 2000), *Interaction with Disabled Persons Scale (IDP)* (Gething, 1994) and *Demographic questions.*

Further information about the SIMS, IDP and SRA can be found in Chapter 2, section 2.8.

8.2.2.4 TagIT project

TagIT is a crowdsourcing project to generate *tags* of image. In this project, *tags* refer to a word or several words used to identify an object in an image. For example, a set of tags suitable for the slide presented in Figure 8.1 might be: Magazine cover, black African boy, black beret with red trim, khaki jacket, green shirt.

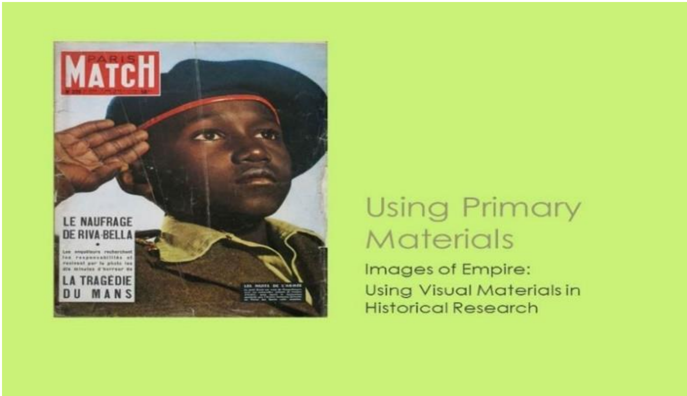


Figure 8.1An example slide used in teaching materials

For the TagIT project, participants were asked to tag individual objects within images of PowerPoint slides from Professor Helen Petrie's inaugural lecture, titled: "Navigating in the dark: Technology for disabled and elderly people". This lecture was the one used in Studies 6 and 7.

To help students create good image tags, the project information page provided tips on how to create tags and example tags of a typical image (see Appendix E.1.2). The structure of the project was the same as the DescribeIT project (see Chapter 3 section 3.2), except that students in the TagIT project were asked to tag the images instead of describing them. This structure was similar to Feyisetan et al. (2015), in which the "traditional crowdsourcing technique" condition, used an "image field – text field" layout to label images. Similar to the DescribeIT project, once students decide to contribute to the project they were presented with a slide and a text-box in which to type their tags of the image on the slide. Figure 8.2 illustrates the project layout. Students were told they could tag as much images as they wished in as many sessions as they wished.

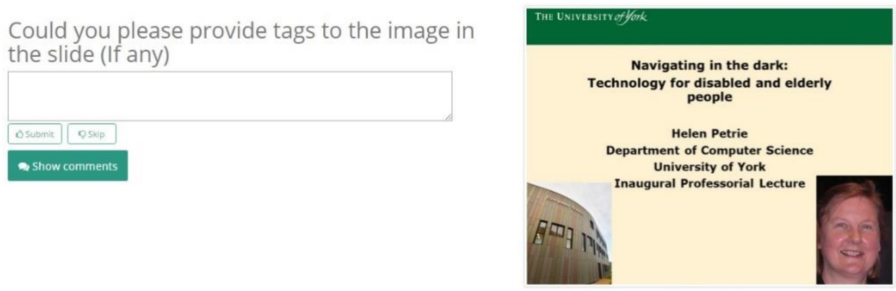


Figure 8.2 The layout of TagIT project

8.2.2.5 Data Analysis: Tag quality

To assess the quality of the tags produced in this study, four PhD students were asked to assess if the tags were related to the content of the images. Two students assessed the same set of tags, and they were asked to:

1. Decide if the tags are relevant to the image (Yes, they are relevant, No, they are not relevant).

2. Rate to what extent the tags are relevant to the image (1 = not at all relevant, 5 = very relevant to).

3. Explain their decision. In case there was a tag that is not relevant to the image it can be mentioned here.

A tag was considered to be appropriate if two or more students agreed that it did represent an object in the image. This approach was based on the ESP game (von Ahn and Dabbish 2004), where two independent players take turns to name objects from the same image, and if the two players tags match, they earn points. As noted by Kennedy, Slaney, and Weinberger (2009) the game "relies on the assumption that agreement between two independent human annotators is sufficient for determining the reliability of a given annotation" (p18).

8.2.2.6 Pilot Study

A pilot study with two undergraduate students revealed some minor design flaws in the TagIt project, which were resolved. These pilot study participants were excluded from the main study.

8.2.2.7 Procedure

Separate recruitment emails were sent out to the undergraduate and the MSc students inviting them to participate in the TagIT project (see Appendix E.1.1). A reminder email sent out to each group one week after the first recruitment email.

8.2.2.8 Data Analysis

The variables in the data collected were normally distributed, except for the SIMS sub-scale External Regulation. However, to make sure that the data is comparable across all the studies, nonparametric tests were used throughout the analysis. Equivalent parametric tests (one-sample t-test, Pearson correlations) showed the same results, except for SIMS sub-scale, Identified Regulation, which will be discussed with caution.

8.2.3 Results

24 students participated in the TagIT project to tag objects in the images, a response rate of 24%. Students tagged a total of 174 images, a mean of 7.3 (SD=7.9) images per student.

The mean and median ratings and standard deviations of students' scores on the four sub-scales of the SIMS, the two sub-scales of the IDP, and their overall SRA scores are summarized in Table 8.1.

Table 8.1 Means, and median ratings and standard deviations and significant tests for SIMS, IDP, and SRA scales

Measure	***TagIT Project (N=14)***			
	Mean (SD)	**Median**	**W**	***P***
SIMS Subscales				
Intrinsic Motivation	4.4 (1.1)	4.7	1.6	n.s
Identified Regulation	4.2 (1.2)	4.0	0.6	n.s
External Regulation	2.5 (0.9)	2.3	-3.1	0.002
Amotivation	2.2 (0.9)	2.0	-3.2	0.001
IDP Scale	75.0 (6.8)	76.0	3.1	0.002
IDP Sub-scales				
Sympathy	25.6 (2.6)	27.0		
Discomfort	13.1 (4.4)	12.0		
SRA Scale	41.4 (7.9)	39.2	-1.5	n.s

A one-sample Wilcoxon signed rank test was conducted to investigate whether the scores on each of the SIMS sub-scales were above the midpoint of 4 or not. Table 8.1 shows that for the "Intrinsic Motivation", and "Identified regulation" sub-scales the scores were not significantly different from the midpoint. However, the scores for the "External Regulation" and "Amotivation" sub-scales were significantly below midpoint of the scale.

The mean IDP score was 75.0 (SD= 6.8) and a median of 76. A comparison with Brown et al. (2009) reported results on the IDP for British undergraduate students was made, in which their mean IDP score was 62.31. A one-sample Wilcoxon Signed Rank test found that the scores of the participants was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=3.1$, $p<0.01$). On the IDP sub-scales students scored a mean of 25.6 (SD=2.6) and a median of 27.0 on the Sympathy sub-scale and a mean of 13.1 (SD=4.4) and a median of 12.0 on the Discomfort sub-scale.

On the SRA students scored a mean of 41.4 (SD=7.9) and median of 39.2, with a lowest score of 32 and highest score of 59. A one-sample Wilcoxon Signed Rank test found that the scores of the current group of students is not significantly different from those of scores provided by Rushton et al. (1981).

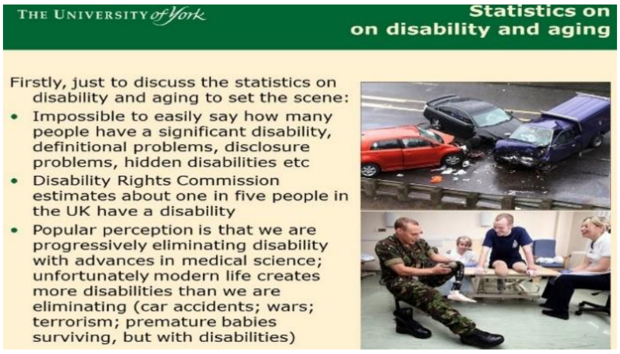
Table 8.2 Correlations between SIMS, IDP, SRA scales and the number of images tagged in the TagIT project.

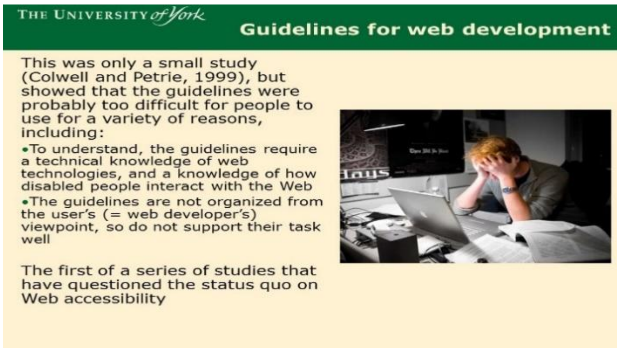
Measures	***p***
SIMS Sub-scales	
Intrinsic Motivation	0.046
Identified Regulation	n.s.
External Regulation	n.s.
Amotivation	n.s.
IDP Sub-scales	
Sympathy	n.s.
Discomfort	n.s.
SRA Scale	n.s.

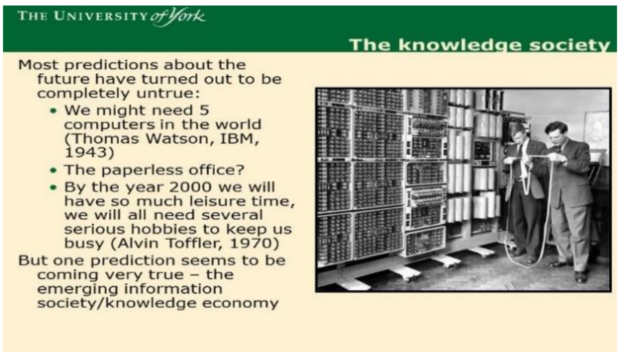
To investigate whether there was a relationship between the number of images tagged and the scores on the SIMS, IDP and SRA, a series of Spearman correlation tests were conducted. These are summarized in Table 8.2. There was significant positive correlations only between the number of images tagged and the Intrinsic Motivation sub-scale of the SIMS: $r_s (14) = 0.54$, $p<0.05$

The quality of 17 (10%) images tagged was assessed. All the tags generated were assessed as being relevant to the images (i.e. two raters agreed they were relevant). Tag quality was assess on a 5-point Likert item (1 = not at all relevant, 5 = very relevant), with an overall mean rating of 4.5 (SD=0.4). Interestingly, students often created not single word tags, but simple descriptions which could easily be built into full descriptions. Table 8.3 shows examples of students' useful tags.

Table 8.3 Examples of students' useful tags

<table>
<tr><th>Images</th><th>Tags</th></tr>
<tr><td>


</td><td>2 colour images
upper: 3 way car crash, all front on
lower: army vetrans [veterans] with amputees receive prosthetic legs from nurses</td></tr>
<tr><td>


</td><td>Colour image - Frustrated man at desk with head in hands and open mac book</td></tr>
<tr><td>


</td><td>A black and white image
two men looking at a wire in a room with old computers circa. 1940-50</td></tr>
</table>

8.2.4 Discussion and Conclusions

This study investigated students' participation in a socially responsible crowdsourcing project aimed to tag individual objects within images for blind and partially sighted students. The tagging task was an attempt to make the description task easier to sighted students in comparison to the task used in the previous studies. Potentially lecturers or other students could build up on these tags to create useful descriptions. For results summary see Tables 8.8, 8.9 and 8.10 end of this chapter.

Overall, a total of 24 students participated to the TagIT project, with a participation rate of 24%. This is a much higher participation rate than of in the other studies with students in the UK, Studies 2, 3, and 4 which used the description task. The mean participation rate in those studies was 3.6%[13], so making the task easier has increased overall participation over six fold. In particular, the participation rate in the present study is substantially higher than the participation rate in the ExMot condition in Study 2 (8%), in which the same extrinsic motivational factor, improving one skills was investigated with the description task.

However, it is interesting that none of the undergraduate students participated in the study, all the 24 participants were MSc. students from the Department of Computer Science. This is in line with the findings of the previous studies. Study 2 involved MSc. students and showed a higher participation rate (11%) than the undergraduate students in Studies 3 (Round 1: 0%, Round 2: 3%) and 4 (0.5%). Even in Study 5 with the Arab students, most of the participants were postgraduate students. Perhaps this is because postgraduate students are more involved in research than undergraduate students, which could have influenced their motivation to take part in a crowdsourcing project to support research. This was actually mentioned by some of the students in Study 5.

This study found that students were neutral in their intrinsic motivation to participate to the TagIT project. This is not in line with the results of the results of Studies 5 and 7, in which participants were intrinsically motivated to participate to the DescribeIT project. However, the present study investigated students' motivation doing a tagging task whereas in Studies 5 and 7 participants were doing a description task. Considering that the description task is a difficult one in comparison to the tagging task, this could mean that only highly intrinsically motivated people participated to the DescribeIT

[13] Study 2: 11%, Study 3, Round 1: 0%, Round 2: 3%; Study 4: 0.5%. Mean = 3.6%.

project, which resulted in weakening of the manipulation of the extrinsic motivational factors, namely the ExMot condition in Studies 5 and 7. Whereas with the tagging task students did not need to be highly intrinsically motivated to participate in a simple task like tagging individual objects in images.

Similar to the results of Studies 5 and 6 participants in the present study did not perceive the task as chosen by themselves, as measured by Identified Regulation. This contrasted with the results of Study 7, in which participants did perceive the task as chosen by themselves. Again, the nature of the description and tagging images task in these studies is different. However, as explained in the Discussion section of Study 6, although participants in Studies 5 and 6 and even in the present study took part voluntarily to the studies, the SIMS sub-scale Identified Regulation showed that students were neutral about their feeling that one has voluntarily participated. Whereas in Study 7 participants were actively participating to crowdsourcing projects.

In line with the findings of Studies 5, 6, and 7, students in the present study were not extrinsically motivated to participate in a socially responsible crowdsourcing project, even when extrinsic motivational factors were highlighted in the studies. It is understandable that people would not like to think that their participation to a project which could potentially help disabled students is regulated by extrinsic motivational factors. As with all the previous studies, participants in this study showed very low Amotivation scores, meaning that students valued their participation in the TagIT project.

The level of participation was also measured by the number of images tagged by each student, and it was hypothesized that there would be a positive relationship between the number of images tagged and the students' sense of altruism, positive attitudes towards disabled people (discomfort and sympathy), and their situational motivation. However, there was no relationship between participation in the TagIT project and any of these variables, except for the Intrinsic Motivation sub-scale of the SIMS scale on which there was a positive correlation between Intrinsic Motivation scores and the number of images tagged.

Lastly, the quality of 10% of the tags was measured as another way of measuring students' level of participation in the task. This showed that students provided a relevant and useful tags. Interestingly, students did not tag image objects with just one or two words, instead they often created simple descriptions that can be easily be b uilt into useful detailed descriptions.

In summary, making the description task simpler task to do could potentially increase students' participation rate to make images of digital learning resources more accessible to blind and partially sighted students. In addition, the tags produced by students were of good quality and useful to build up on them to create descriptions.

8.3 Study 9: Participation in the DescribeIT versus TagIT projects in a public crowdsourcing project

8.3.1 Introduction

To improve the external validity of the results obtained in Study 8, it was decided to conduct Study 9 with crowd members. In addition, Study 9 sought to investigate if the participation rate would be higher from a group of people who are already active members of a crowd, that is who are already actively involved in crowdsourcing tasks. Therefore, the research questions to be addressed in this study were: firstly, would more members of an existing crowd participate in the image tagging task in comparison to the image description task? Secondly, would members of a crowd show higher participation rates than students in both the image tagging and image description tasks? Finally, would the quality of the tags and descriptions produced by active crowd members be as good as those produced by students?

8.3.2 Method

8.3.2.1 Design

The study investigated the participation rates of crowd members from the Crowdcrafting.org platform in two socially responsible crowdsourcing projects: DescribeIT and TagIT projects. A comparison was made between the participation rates in the DescribeIT project which asks participants to describe images and the TagIT project in which participants are asked to tag individual objects within the image. The hypothesis is that the participation rate on the TagIT project will be higher that on the DescribeIT project. In addition, a comparison was made with the participation by students in the same projects (data from Studies 2, 3 and 4 for the image description task and Study 8 for the tagging task). The hypothesis is that the participation by members of the Crowdcrafting platform (known as "crafters") will be higher than the participation rate of students. Finally, a comparison of the tag and descriptions quality was made between the crafters and the students. The hypothesis is that the quality of tags and descriptions from the students will be higher than from the crafters.

The independent variable was the nature of the task, with two conditions: the tagging images task (TagIT project) or the describing images task (DescribeIT project). In both conditions, the extrinsic motivation of improving participants own skills was highlighted on the project information page. The dependent variables were the number of people who participated in each condition, the number of tags produced by each participant (in the tag condition) and the number of descriptions produced by each participant (in the descriptions condition). In addition, the quality of both the image tags and descriptions produced were measured. In line with previous studies other variables were measured using an online questionnaire, including:

1. Participants' motivation while doing the tasks: measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000).

2. Participants' overall attitude toward disability: measured using the Interaction with Disabled Persons Scale (IDP) (Gething, 1994).

3. Participants' overall level of altruism: measured using the Self-Report Altruism Scale (SRA) (Rushton et al. 1981).

8.3.2.2 Participants

A total of 111 participants took part in this study, 79 in the TagIT condition and 32 in the DescribeIT condition. However, only 39 participated in the online questionnaire, including 26 in the TagIT condition and 13 in the DescribeIT condition. This comprised 15 women and 10 men in the TagIT condition (one participant preferred not to give their gender), and 4 women and 4 men in the DescribeIT condition (5 participants did not answer the demographic questions). In the TagIT condition, participants' age range was 27 to 47 years old, with mean age of 34.0 (SD = 5.8) years old. In the DescribeIT condition, the age range was 21 to 34 years old, with mean age of 25.8 (SD = 5.7) years old.

8.3.2.3 Materials

In the TagIT project participants were told: *Participating in this project will also improve your skills in analysing and tagging images which will make your blog articles, YouTube videos and images easy to find by search engines. Following our image tagging tips will help you learn how to provide good image tags.* In the DescribeIT project participants were told: *Participating in this project will improve your skills to make websites, apps and software accessible to people in the future. Following the image* *description guidelines provided with this system will help you learn how to provide good image descriptions.*

As in the previous studies, the online questionnaire of both conditions consisted of four sections (See appendix C.13), including: *Self-Report Altruism Scale (SRA)* (Rushton et.al. 1981), *Situational Motivation Scale (SIMS)* (Guay, Vallerand, and Blanchard, 2000), *Interaction with Disabled Persons Scale (IDP)* (Gething, 1994) and *Demographic questions.*

Further information about the SIMS, IDP and SRA can be found in Chapter 2, section 2.8.

8.3.2.4 TagIT and DescribeIT projects

The TagIT project used in this study was the same one discussed earlier in this chapter, see section 8.2.2.4. The DescribeIT project used in this study was the same one as discussed in Chapter 3, see section 3.2.

In both projects (TagIT and DescribeIT) the same pack of slides was used. The images used in both projects were from the York Digital Library[14]. In total, each project had 32 images. The aim was for each image to be tagged by 30 times (for valid results) in the TagIT condition, which would mean a total of 960 images were available. For the DescribeIT condition the aim was for each image to be described 30 times, which would mean a total of 960 descriptions.

8.3.2.5 Data Analysis: Quality of Tags and Descriptions

To assess the quality of the tags produced in the TagIT project, the same measures used as in Study 8 (see section 8.2.2.5) were used. The participants were the same ones as those who rated tags in Study 8.

The same measures of image description quality were used as in Study 6: word count, SDT method and expert assessment. The SDT method and expert assessment was carried out on 10% of the descriptions produced (10 descriptions). The experts were the same ones who participated in Study 6 (See section 7.2.1.5 for more details).

8.3.2.6 Pilot Study

A pilot study with two research students was conducted. Each participant was asked to contribute to either the DescribeIT or TagIT project, and respond to the online

[14] https://dlib.york.ac.uk/yodl/app/home/index

questionnaire. No technical problems were detected. Neither pilot participant took part in the main study.

8.3.2.7 Procedure

The projects were created on the Crowdcrafting.org platform. Both the TagIT and DescribeIT projects were published on the same day at the same time on the platform. The reasons for publishing the two projects at the same time was, firstly, to measure participants' propensity to choose one or other of the projects based on the nature of the project, which might be affected if the two projects were launched at different times. Secondly, it was to make sure that both projects were affected equally by the same availability of other projects available on Crowdcrafting.org at the same time, which the researcher might not be aware of and cannot control.

Both projects were featured by the Crowdcrafting support team. Featured projects (See Figure 8.3) are posted on the Crowdcrafting.org home page.

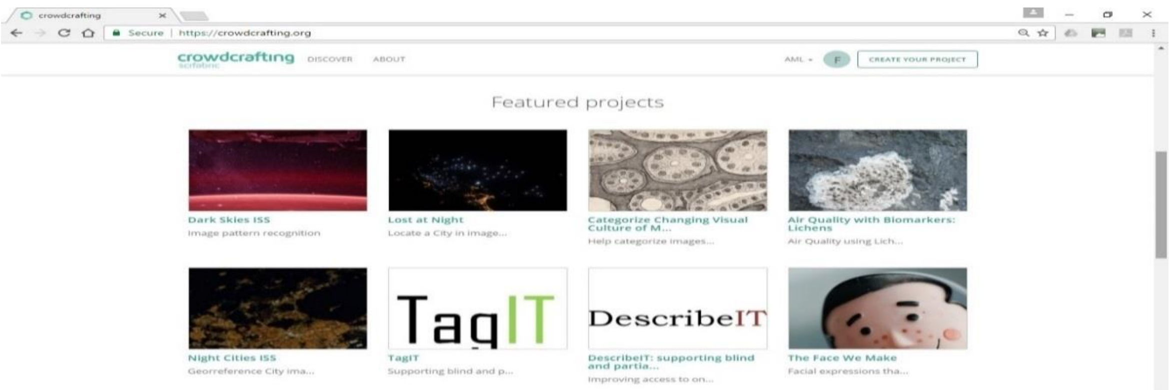


Figure 8.3 Screenshot of Crowdcrafting.org home page

In addition, as shown in Figure 8.4 (as an example) Professor Petrie tweeted about both projects to recruit participants. Two fellow researchers retweeted her tweets to recruit participants.

Figure 8.4 Professor Petrie's tweet about the DescribeIT project

8.3.2.8 Data Analysis

As with previous studies, nonparametric tests were used throughout the analysis of both conditions. The equivalent parametric tests one-sample t-test, and Pearson correlations showed the same results when conducted with normally distributed variables.

8.3.3 Results

32 crafters participated in the DescribeIT project condition. Participants viewed 174 images and produced a total of 102 image descriptions, meaning they skipped 72 images. They described a mean of 4.3 (SD=6.9). The 13 crafters who participated in the online questionnaire produced a total of 67 image descriptions, a mean of 5.2 (SD=2.1) descriptions per crafter. 79 crafters participated in the TagIT project condition. These participants viewed 486 images, tagging a total of 312 images, meaning they skipped 174 images. They tagged a mean of 5.6 (SD=1.7) images per crafter. The 26 crafters who participated in the online questionnaire tagged a total of 156 images, a mean of 6.0 (SD=1.9) images per crafter. A Mann-Whitney test indicated no significant differences between the number of images tagged and described (U=762, p=0.326).

The mean and median ratings and standard deviations of participants' scores on the four sub-scales of the SIMS, the two sub-scales of the IDP, and their overall SRA scores for the DescribeIT and TagIT projects are summarized in Table 8.4 and 8.5 respectively.

One-sample Wilcoxon signed rank tests were conducted to investigate whether the scores on each of the SMIS sub-scales were above the midpoint of 4. Table 8.4 shows that in the DescribeIT condition the scores for "Intrinsic Motivation" and "Identified regulation" sub-scales were significantly above the midpoint of the scale. Whereas "External Regulation" and "Amotivation" sub-scales were significantly below midpoint of the scale. Similarly, Table 8.5 shows for the TagIT project the scores for the "Intrinsic Motivation", and "Identified regulation" sub-scales were significantly above the midpoint. However, the scores for the "External Regulation" and "Amotivation" sub-scale were significantly below midpoint of the scale.

For the DescribeIT project condition participants showed a mean SRA score of 44.3 (SD=7.5), with a lowest score of 32 and highest score of 59. A one-sample Wilcoxon Signed Rank test found that the scores of the participants is not significantly different from those of scores provided by Rushton et al. (1981). For the TagIT project condition participants showed a mean SRA score of 53.6 (SD=8.9), with a lowest score of 41 and highest score of 77. A one-sample Wilcoxon Signed Rank test found that the scores of participants was significantly higher than the median of SRA scores provided by Rushton et al. (1981) (W=4.1, p<0.001).

Table 8.4 Mean and median scores, standard deviations and significance tests for SIMS, IDP, SRA scales in the DescribeIT Project

Measures	***DescribeIT Project (N=13)***			
	Mean (SD)	**Median**	**W**	***P***
SIMS Sub-scales				
Intrinsic Motivation	4.9 (1.0)	4.8	1.29	0.003
Identified Regulation	4.5 (1.1)	4.3	2.00	0.045
External Regulation	2.5 (1.1)	2.3	-2.84	0.004
Amotivation	2.2 (1.1)	2.0	-2.94	0.003
IDP Scale	69.0 (7.8)	63.0	2.6	0.008
IDP Sub-scales				
Sympathy	23.2 (3.2)	23.2		
Discomfort	13.4 (4.4)	13.4		
SRA Scale	44.3 (7.5)	44.3	-0.2	n.s

Table 8.5 Mean and median scores, standard deviations and significance tests for SIMS, IDP, SRA scales in the TagIT project

Measures	***TagIT project (N=26)***			
	Mean (SD)	**Median**	**W**	***P***
SIMS Sub-scales				
Intrinsic Motivation	5.1 (1.1)	5.0	3.82	0.000
Identified Regulation	4.8 (0.8)	5.0	3.64	0.000
External Regulation	2.2 (1.5)	2.0	-3.97	0.000
Amotivation	2.2 (1.6)	2.0	-3.78	0.000
IDP Scale	67.0 (12.5)	66.5	2.0	0.049
IDP Sub-scales				
Sympathy	22.9 (3.7)	23.0		
Discomfort	11.4 (6.2)	10.0		
SRA Scale	53.6 (8.9)	51.0	4.1	0.000

For the DescribeIT project condition, the mean IDP score was 69.0 (SD= 7.8). A comparison with Brown et al. (2009) reported results on the IDP for British undergraduate students was made. A one-sample Wilcoxon Signed Rank test found that the scores of the participants was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=2.6$, $p<0.001$). On the IDP sub-scales participants' scores on the Sympathy sub-scale were a mean of 23.2 (SD=3.2), and on the Discomfort sub-scale a mean of 13.4 (SD=4.4). For the TagIT project condition, the mean IDP score was 67.0 (SD= 12.5). A comparison with Brown et al. (2009) reported results was made. A one-sample Wilcoxon Signed Rank test found that the scores of the participants was significantly higher than the median of IDP scores provided by Brown et al. (2009) ($W=2.0$, $p<0.05$). On the IPD sub-scales participants' scores on the Sympathy sub-scale were a mean of 11.4 (SD=6.2), and 22.9 (SD=3.7) for the Discomfort sub-scale.

Table 8.6 Correlations between SIMS, IDP, SRA scales and the number of images described in the DescribeIT project condition and the number of images tagged in the TagIT project condition

Measure	**DescribeIT**	**TagIT**
SIMS subscale		
Intrinsic Motivation	n.s.	n.s.
Identified Regulation	0.039	n.s.
External Regulation	n.s.	n.s.
Amotivation	n.s.	n.s.
IDP Scale		
Sympathy	n.s.	n.s.
Discomfort	n.s.	n.s.
SRA Scale	n.s.	n.s.

To investigate whether there was a relationship between the number of descriptions produced and the scores on the SIMS, IDP and SRA, a series of Spearman correlation tests were conducted. These are summarized in Table 8.6. For the DescribeIT project there was a significant positive correlation between the number of images described and Identified Regulation sub-scale $r_s(13)=0.58$, $p<0.05$. For the TagIT project there were no significant correlations between the number of images tagged and any of the SIMS sub-scales, IDP sub-scales or the SRA scale.

The quality of 33 (slightly more than 10%) images tagged in the TagIT project was assessed. All tags were assessed as relevant to the images. The quality of the tags was assessed on a 5-point Likert item (1=Not at all relevant, 5=Very relevant), with a mean rating of 3.8 (SD=0.7). This mean rating can be explained by the fact that raters thought that participants sometimes missed objects in the images that were important and because they included the text from the slide as tags. For example, raters commented:

too simple tags! A lot more can be explained [Rater 2]

contain text descriptions [Rater 4]

It is worth noting that although raters were asked to rate to what extend the tags were relevant to the images, it is clear from their comments they thought of the overall quality of the tags, not only how relevant the tags are.

Table 8.7 illustrates the frequency and percentage of the word count range in the image descriptions produced in the DescribeIT project. Overall the mean number of words per description was 18.2 (SD= 18.3, range 1 to 95 words).

Table 8.7 The Frequency and Percentage of the Word Count Ranges in the Image Descriptions

Words Range	**DescribeIT**	
	Frequency	Percentage
1 to 25 words	74	73%
26 to 50 words	22	22%
51 to 75 words	4	4%
76 to 100+ words	2	2%
Total	**102**	**100%**

In addition, the quality of 10 (10%) of the image descriptions produced in the DescribeIT project was assessed. The scores of SDT formulae for each of accuracy, error score, and efficiency are summarized in Table 8.8.

In the expert assessment, experts rated the quality of the descriptions on a 5-point Likert item (1=Very poor, 5=Excellent), this gave a mean of 3.6 (SD=0.6). A one-sample Wilcoxon signed rank test was conducted to investigate whether the experts' ratings were above the midpoint of 3 or not. The results showed that experts' ratings were significantly above the neutral midpoint of the scale ($Z=2.3$, $p<0.05$).

Table 8.8 Mean and median scores and standard deviation of accuracy, error score, and efficiency of the image descriptions produced in the DescribeIT project

SDT Measure	**Mean**	**Median**	**SD**
Accuracy	58.5	59.1	17.9
Error score	0.0	0.0	0.0
Efficiency	21.6	22.2	6.8

A Spearman rank-order correlation test was conducted to determine if there were any relationships between accuracy level and experts' ratings. The results showed that there was a significant positive correlation between the accuracy and experts' ratings, r_s (10) =0.64, p <0.05.

Experts also decided whether an image description was suitable as a short or long description (see section 7.2.1.5 for further details). A Mann-Whitney U test compared the experts' decisions with the efficiency scores of the descriptions. It showed no significant different in level of efficiency between the descriptions deemed suitable as short and long. However, in two descriptions out of the ten which were assessed, two experts gave different decisions on the description type with the same level of confidence of their decision. These two descriptions were not included in the Mann-Whitney test, this means only eight descriptions were included in the analysis.

8.3.4 Discussion and Conclusions

This study investigated active crowd members participation in two socially responsible crowdsourcing projects: the TagIT project which asked crowd members to tag objects in images for blind and partially sighted students, and the DescribeIT project which asked crowd members to describe images for blind and partially sighted students. For results summary see Tables 8.9, 8.10 and 8.11 end of this chapter.

The level of participation of each project was first measured by the number of participants in each project, and it was hypothesized that the number of participants in the TagIT project will be higher than the number of participants of the DescribeIT project. Overall, 32 crafters participated in the DescribeIT project in comparison to 79 crafters in the TagIT project. The number of crafters who participated in the tagging task was nearly 2.5 times the number who participated in the description task, so the first hypothesis is upheld, making the task easier which clearly encourages more people to participate. Similar findings were reported by Yuen, King, and Leung (2012) who investigated the criteria of task selection by 100 Amazon MTurk workers via an online questionnaire. Their results showed that 86% of the MTurk workers selected their tasks based on the nature of the task, where 58% selected their tasks based on the level of difficulty of the task.

Addressing the second hypothesis, do crafters participate more than students, is more difficult in spite of the wealth of data collected. As the total number of crafters who viewed the project descriptions on the platform website is not known and a snowball recruitment method was used, it was impossible to calculate the participation rate in each project to compare to the participation rate of students. Crowdcrafting has nearly 15,000 registered volunteers[15], so as a proportion of the total possible population of volunteers, 32 and 79 participants is hardly high. Even if one estimates that 5% of volunteers saw a particular project, this would yield participation rates of 4.2% for the DescribeIT Project (32/750) and 10.5% for the TagIT Project. The participation rate in the DescribeIT project is then quite comparable with the overall participation rate by students (in Studies 2, 3 and 4) of 3.4% and the participation rate in TagIT is actually lower than the participation rate by students of 24% (see Study 8). However, these comparisons need to be treated very cautiously as the estimates for Crowdcrafting are very speculative.

The level of participation of each project was also measured by the number of images described or tagged by each participant. However, there was no significant difference between the number of images tagged and those described. Thus, making the task easily increased participation rate, but it did not affect the number of images tagged or described.

Lastly, the quality of the tags and descriptions produced was investigated. Similar to the results of Study 8, all the tags produced in the TagIT project were assessed as relevant to the image content. However, the raters thought that crafters had missed out some important tags. Nonetheless, although the instructions asked crafters to add as much tags as they wished as long as they are relevant to the content of the image, they were not explicitly told to tag every single object in the image. So, missing important objects to tag cannot be considered a good measure of lack of participation and engagement with the task. In any case, in a practical application of the TagIT project, a

15 https://crowdcrafting.org/account/

lecturer or student could aggregate tags produced for a particular image by different crafters (or students) to produce a complete description, so individual participants missing tagging particular objects would not be a problem.

The accuracy of the descriptions produced in the DescribeIT project in the present study was similar to those obtained in Studies 6 and 7. Indeed, the efficiency level of the descriptions in the present study was similar to those obtained in Study 6, but higher than the efficiency level of descriptions in Study 7.

It was hypothesized that there would be a positive relationship between the number of images described or tagged and crafters' sense of altruism, attitudes towards disabled people (both discomfort and sympathy), and their situational motivation. However, there was no relationship between participation in the DescribeIT project and TagIT project and any of these variables, except for the Identified Regulation sub-scale of the SIMS scale on the DescribeIT project, with a positive correlation between the number of images described and the Identified Regulation sub-scale scores.

The DescribeIT project results showed that participants were intrinsically motivated to participate to the project, as measured by the SIMS sub-scale, Intrinsic Motivation. This is in line with the results of Studies 5 and 7, in which participants were significantly intrinsically motivated to participate to the DescribeIT project. Considering the following results: Studies 2, 3, and 4 in which UK students' participation rates were extremely low (despite the reminder emails and second rounds). The results of Study 6 (UK students only participated after been offered money) students were significantly below the midpoint on intrinsic motivation. in Studies 5 and 7significantly above the midpoint on intrinsic motivation and participation was substantially higher. All these results suggest that for people to do the image description task they need to be highly intrinsically motivated. This particular finding suggests another research question: How do participants perceive the difficulty level of the image description task? To understand if students' lack of participation to the DescribeIT project was due to how they think the image description task is difficult, the next study (Study 10) will investigate students' perception of the level of the difficulty of the description task before doing the task and their perceptions after having done the task.

Whilst students in Study 8 were neutral in their intrinsic motivation to participate to the TagIT project, participants in the present study were significantly intrinsically motivated to participate in the TagIT project. This finding agrees with the findings about the DescribeIT project in studies 5, 7, and the DescribeIT condition in the present study.

As in Study 7, participants in the present study in both the DescribeIT and TagIT projects perceived the task as chosen by themselves, as measured by the SIMS sub-scale Identified Regulation. This contrasted with the results of Studies 5, 6, and 8, in which students did not perceive the task as chosen by themselves, although they voluntarily participated. However, this finding is in agreement with the suggestion that members of the crowd in both Crowdcrafting and Amazon MTurk perceive the task as being chosen by themselves in a sense of one has voluntarily chosen to participate to the DescribeIT project. Whereas students in Studies 5, 6, and 8 did not because they did not actively seek to participate to crowdsourcing.

In line with the findings of Studies 5 to 8, participants in the present study in both conditions were not extrinsically motivated to participate. This perhaps because participants' intrinsic motivation overcame their extrinsic motivation to participate in a crowdsourcing project to support disabled students. However, the results of Study 6 do not support this argument, as students for that study were recruited with different intrinsic and non-financial extrinsic motivational factors, but only participated when they were offered money. These finding suggest the influence of social desirability bias as an explanation. Further research is needed to further investigate whether the SIMS scale is subject to social desirability bias, especially with socially responsible tasks.

Participants of both the DescriebIT and TagIT projects showed a very low Amotivation scores, which is understandable considering that the participants of both projects were significantly intrinsically motivated to participate in the projects. The fact that this finding was also found in previous studies (Studies 5 to 8) means that participants in these studies valued the project to some degree or the outcomes of it.

To sum up, the findings of this study are in agreement with the findings of Study 8, in which making the task simpler increased the participation rate in a socially responsible crowdsourcing task. The quality of the tags and descriptions produced were similar to those produced by students. Participants of both projects were significantly intrinsically motivated to participate.

Table 8.9 Summary of the results of Studies 2 to 9

Measures	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7	Study 8	Study 9 TagIT	Study 9 DescrbeIT
Number of participants (phase 1)	27 students	26 students	36 students	92 students	26 students	Literature Review	-	-	-
Number of participants (phase 2)	8 students	No students	2 students	89 students	26 students	164 workers	24 students	79 Crafters	32 Crafters
Number of participants (phase 2 online questionnaire)	No students	phase 3: 19 students	1 student	67 students	14 students	86 workers	14 students	26 Crafters	13 Crafters
Number of descriptions	86 descriptions	17 descriptions	11 descriptions	584 descriptions	385 descriptions	851 descriptions	174 images tagged	312 images tagged	102 descriptions

Table 8.10 Results of One-sample Wilcoxon tests for the SIMS scale of Studies 5 to 9

SIMS Scale	Study 5 (InMot)	Study 5 (ExMot)	Study 6	Study 7	Study 8	Study 9 TagIT	Study 9 DescrbeIT
Intrinsic Motivation	+**	+**	-*	+**	n.s	+**	+**
Identified Regulation	n.s	n.s	n.s	+**	n.s	+**	+*
External Regulation	-**	-*	-**	-**	-**	-**	-**
Amotivation	-**	-**	-**	-**	-**	-**	-**

Table 8.11 Correlation between number of images and SIMS, IDP, and SRA scale of Studies 5 to 9

SIMS Scale	**Study 5 (InMot)**	**Study 5 (ExMot)**	**Study 6**	**Study 7**	**Study 8**	**Study 9 TagIT**	**Study 9 DescrbeIT**
Intrinsic Motivation	n.s	n.s	n.s	n.s	+*	n.s	n.s.
Identified Regulation	n.s	+**	n.s	n.s	n.s	n.s	+*
External Regulation	n.s	+*	n.s	n.s	n.s	n.s	n.s
Amotivation	n.s	-*	n.s	n.s	n.s	n.s	n.s
IDP Scale							
Sympathy	n.s	+**	n.s	n.s	n.s	n.s	n.s
Discomfort	n.s	n.s	-*	n.s	n.s	n.s	n.s
SRA Scale	n.s	+*	n.s	n.s	n.s	n.s	n.s

Chapter 9

Training and Quality Control to Improve Engagement and Performance with a Socially Responsible Crowdsourcing Project

9.1 Introduction

The results of the previous Studies 2 to 9 suggest that the task of describing images for blind and partially sighted people is difficult for crowd members. This is probably in part because the task requires a good and confident understanding of what to describe and what not to describe in an image, what to include and what not to include in a description. In all the previous studies, participants had no experience of the image description task or of supporting blind and partially sighted students; crowd members were simply students or members of a crowdsourcing platform who volunteered to take part. The only information they had was a set of guidelines to describe images for blind and partially sighted students and an example description. However, considering the previous results in terms of both number of participants who volunteered and the quality of the image descriptions produced, simply presenting potential participants with a set of guidelines may not be sufficient to allow them to acquire the necessary understanding to provide appropriate descriptions.

This chapter will present the last study *(Study 10)* in this programme of research which investigates whether providing students with some face-to-face training increases their participation in the DescribeIT Project and whether it increases the quality of the descriptions produced. The question of participants' perception of the difficulty of the image description task will also be investigated. Finally, this study investigates the effect of different quality control messages in the instructions about the task on the quantity and quality of descriptions produced.

This study sought to answer the following research questions:

- Does a face-to-face training session improve students' motivation to participate in the image description task and improve the quantity and quality of the image descriptions produced?
- Does a face-to-face training session decrease students' perception of the difficulty of the image description task?
- Do instructions about quality control (i.e. that a peer or an expert will check the quality of the descriptions) increase the quality of the descriptions produced?
- Do attitudes to people with disabilities correlate with the quantity and quality of the descriptions produced?
- Do levels of altruism correlate with the quantity and quality of the descriptions produced?

9.2 Method

9.2.1. Design

This study investigated the effect of a face-to-face training session on students' motivation to participate in the image description task and the quantity and quality of the images produced. It also investigated the effects of quality control messages in the instructions for the image description task. It also investigated students' perceptions of the difficulty of the image description task and whether this was altered by the training.

The study consisted of three phases, as follows (see Figure 9.1, below):

Pre-Study Phase: an online questionnaire in which students rated the difficulty of the image description task (Initial Difficulty Rating) after they had read an explanation of the task. This questionnaire also collected demographic data, including students' gender, age, nationality, previous crowdsourcing experience [See Appendix F.1.1].

Pre-Training Phase: Students were randomly assigned to one of the three Quality Control conditions. They were asked to participate in the DescribeIT project for 20 minutes. They then rated the difficulty of the image description task again (PreTraining Difficulty Rating) and completed the Situational Motivation Scale (SIMS) to assess their motivation to participate in the task [See Appendix F.1.2].

Training: Students were then given 20 minutes training in how to undertake the image description task by an expert in the field (Professor Helen Petrie).

Post-Training Phase: Students were asked to participant in the DescribeIT Project for a further 20 minutes. They then rated the difficulty of the image description task again (PostTraining Difficulty Rating) and completed the *Situational Motivation Scale (SIMS)* (Guay, Vallerand, and Blanchard, 2000) again to assess their motivation to participate in the task after training. They also completed the *Interaction with Disabled Persons Scale (IDP)* (Gething, 1994) to assess their attitude to people with disabilities, and the *Self-Report Altruism Scale (SRA)* (Rushton et.al. 1981) to assess their level of altruism [See Appendix F.1.3].

Thus, the study had a two-way, mixed design with one between-participant independent variable and one within-participant independent variable. The between-participant independent variable was the Quality Control information in the instructions, which had three levels: no quality control (NoQC), quality control conducted by other students participating in DescribeIT (PeerQC), or quality control conducted by an expert (ExpertQC). The within-participant independent variable was Training, with two levels (Training or No Training).

The dependent variable related to students' motivation to do the image description task was their scores on the SIMS, taken after the PreTraining experience with the task and again after the PostTraining experience with the task.

The dependent variable related to the image description task were the number of images described by each student and the quality of the image descriptions (see section 9.2.5 for details of how image quality was measured).

The dependent variable related to the students' perception of the difficulty of the image description task were their ratings of task difficulty taken before they had experienced the task (but had read a description of it) (Initial Difficulty Rating), after they had experienced the task without training (although they had received the guidelines and an example) (PreTraining Difficulty Rating) and after they had experienced the task after training (PostTraining Difficulty Rating).

The dependent variable relating to students' attitudes to people with disabilities was their scores on the Interaction with Disabled Persons Scale (IDP) (Gething, 1994).

The dependent variable relating to students' altruism was their scores on the Self-Report Altruism Scale (SRA) (Rushton et.al. 1981).

Further information about the SIMS, IDP and SRA can be found in Chapter 2, section 2.8.

Pre-Study Phase		
Pre-Study questionnaire: Verbal explanation of the DescribeIT Project and the image description task Initial Task Difficulty Rating (Initial Difficulty Rating) Demographic information		
Pre-Training Phase		
Quality Control Conditions		
No Quality Control (NoQC)	Peer Quality Control (Peer QC)	Expert Quality Control (Expert QC)
Instructions: written guidelines and example of image description task (with appropriate Quality Control information) DescribeIT image description task (20 minutes)		
Pre-Training questionnaire: Pre-Training Task Difficulty Rating (PreTraining Difficulty Rating) Situational Motivation Scale (SIMS)		
Face-to-face training in Image Description for blind and partially sighted students by an expert (20 minutes)		
Post-Training Phase		
DescribeIT image description task		
Post-Training questionnaire: Post-Training Task Difficulty Rating (PostTraining Difficulty Rating) Situational Motivation Scale (SIMS) Interaction with Disabled Persons Scale *(*IDP) Self-Report Altruism Scale (SRA)		

Figure 9.1 Overall design of Study 10

9.2.2. Participants

140 undergraduate students were invited to take part in this study as part of the practical work for their module on Human Aspects of Computer Science. Students were undergraduate students from the Department of Computer Science at the University of York. 128 students chose to participate in the study, comprising 11 women and 116 men (one student preferred not to give their gender). Their age range was 17 to 26 years old, with mean age of 18.5 (SD = 1.0) years. The majority of students were from United Kingdom (107, 84 %), 4 (3%) students were from China, 2 (2%) students were

from France, and the remaining students (11%) were spread relatively evenly across 14 other countries: Greece, Slovakia, Sweden, Kenya, Poland, UAE, Bangladesh, USA, Bulgaria, Romania, Italy, Lithuania, South Africa, and Gibraltar. One student preferred not to say where they were from.

All the students reported using social media such as Facebook, Instagram and Snapchat. However, only 25 students (19.5%) reported having participated in crowdsourcing projects, including: Galaxy Zoo, Zooinverse, Wikipedia and many others.

Students were randomly assigned to one of the three conditions of Quality Control (NoQC=47 students, PeerQC=47, and ExpertQC=46). However, not all students participated in the study and some of them only participated in some phases. The breakdown of number of students in each condition in all the study phases is shown in Table 9.1. For the online questionnaires, students who did not complete all the questionnaires sections were removed.

Table 9.1 The breakdown of number of students in each condition in each phase

The study phases	NoQC	PeerQC	ExpertQC	Total
DescribeIT Pre-Training	45	38	42	125
Pre-Training Questionnaire	34	30	35	99
DescribeIT Post-Training	41	38	39	118
Post –Training Questionnaire	28	31	31	90

9.2.3. Materials

The **Pre-Study Questionnaire** consisted of three sections:

- *Explanation of the Image Description Task*: gave a general information about the DescribeIt project and its importance for blind and partially sighted students, the image description task and how it would be undertaken in the DescribeIt project, and the type of images to be described.
- *PreStudy Task Difficulty Rating:* a 7-point Likert item of how difficult the image description task would be (1= Extremely Difficult, 7= Extremely Easy)

- *Demographic Information:* information including students' age, gender, and previous experience with crowdsourcing.

Six versions of the DescribeIT Project were created, three for the Pre-Training Phase and three for the Post-Training phase. The three versions for each Phase contained the different information about the Quality Control conditions, as follows:

- NoQC: *Please be aware that no one will check the quality of your descriptions so please provide good descriptions for the blind and partially sighted students.*
- PeerQC: *Please be aware that other participants will check the quality of your descriptions. You will get a feedback on some of your descriptions within a couple of weeks.*
- ExpertQC: *Please be aware that an expert will check the quality of your descriptions. You will get a feedback on some of your descriptions within a couple of weeks.*

Further information about the DescribeIT Project are provided in Section 9.2.4, below.

The **Pre-Training Questionnaire** consisted of two sections:

- *PreTraining Task Difficulty level:* a 7-point Likert item of how difficult the image description task was in the pre-training phase (1= Extremely Difficult, 7= Extremely Easy)
- *Situational Motivation Scale (SIMS):* A 16-item scale to assess motivation and engagement in the image description activity, on 7 point Likert items (1=not at all, and 7= exactly). This includes four subscales: Intrinsic Motivation, Identified Regulation, External Regulation, and Amotivation.

The training session covered the types of images that need descriptions, the aspects of the images that need descriptions, and included examples of good and bad descriptions. A recording of the training session is included on the CD accompanying this thesis.

The **Post –Training Questionnaire** consisted of four sections:

PostTraining Task Difficulty level: a 7-point Likert item of how difficult the image description task was in the post-training phase (1= Extremely Difficult, 7= Extremely Easy)

Situational Motivation Scale (SIMS) (Guay, Vallerand, and Blanchard, 2000),

Interaction with Disabled Persons Scale (IDP) (Gething, 1994)

Self-Report Altruism Scale (SRA) (Rushton et.al. 1981)

For further details about the SIMS, IDP, and SRA, see Chapter 2, section 2.8.

9.2.4. DescribeIT project

The DescribeIT project used in this study was the same as the one presented in Chapter 3, section 3.2. The project information page provided students with instructions and information about the quality control they would get about their descriptions (see Materials, section 9.2.3, above).

The project information page provided guidelines on how to describe images for blind and partially sighted people (the same guidelines as used in all previous studies) and an example description of a typical image. The instructions also indicated that students could describe as many or as few images as they wished.

Once students decided to start the task, they were presented with a slide and a text-box in which to type their description of the image on the slide. Figure 9.2 illustrates the layout of the image description page.

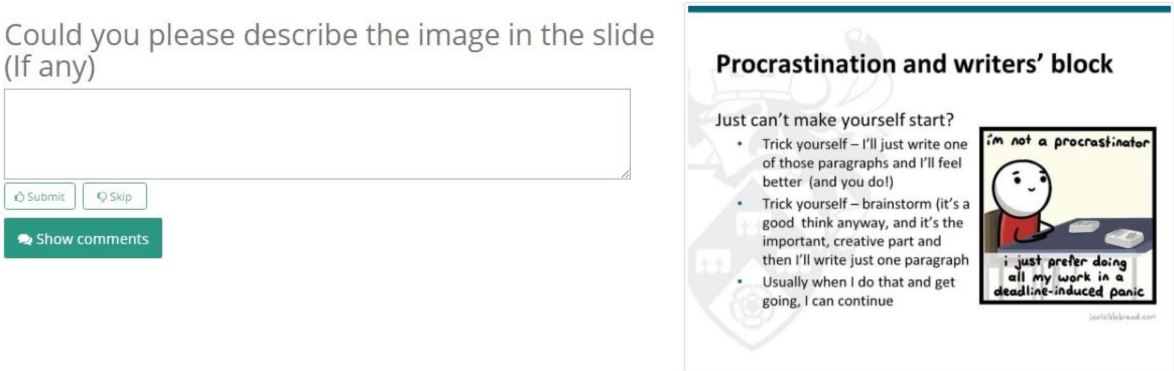


Figure 9.2 The layout of the DescribeIT project

In the PreTraining phase, a set of images from a PowerPoint pack from the module "Skills, Knowledge and Independent Learning" were used. In total, this pack had 33 slides, with 13 images. The aim was for each image to be described by each student in each QC condition, which would mean a total of approximately 611 descriptions (13 images x 47 students) would be generated in each condition.

In the PostTraining Phase, a set of images from another PowerPoint pack from the same module was used. In total, this pack had 28 slides with 13 images. Again, the aim was for each image to be described by each student in each QC conditions, which would mean a total of approximately 611 descriptions in each condition.

9.2.5. Data analysis: Image description quality

The same measures of image description quality were used as in Study 6: word count, SDT method and expert assessment. For the SDT method and expert assessment a total of 96 descriptions was assessed in the PreTraining phase and 96 descriptions was assessed in the PostTraining phase. For the expert assessment, in the PreTraining phase, each expert evaluated 24 descriptions of eight images. There were three descriptions of each image, one from each of the QC condition. Similarly, in the PostTraining phase each expert evaluated 24 descriptions of eight images. There were three descriptions of each image, one from each of the QC condition. The experts were the same one participated in Study 6.

9.2.6. Pilot study

Three PhD students from the University of York piloted different versions of the DescribeIT projects. They were each asked to try one version of the PreTraining and PostTraining Projects (thus one student tried the NoQC version, one tried the PeerQC version and one tried the ExpertQC version) and to respond to the PreStudy, PreTraining and PostTraining questionnaires. No technical problems were detected in the projects. The pilot participants did not take part in the main study.

9.2.7. Procedure

The study was run as part of a practical session in the module "Human Aspects of Computer Science". Students were expected to attend a two-hour session to learn about how to be part of a study. However, students were informed that participation in this study was voluntary and their data would be treated confidentially and anonymously. The great majority of students decided to participate in the study.

The study involved three main phases, all were carried out during the practical session itself. The sequence of activities was as follows:

As students entered the practical laboratory they were randomly assigned to one of the three different QC conditions. Students in each condition were seated in groups and separated from students in the other conditions so they could not see each others instructions or discuss them.

Each student was provided with a hard copy of the study information document, consent form, and a copy of the project information page (see Appendix F.1.4 and

F.1.5). Students were given a verbal introduction about the DescribeIT project and its importance for blind and partially sighted students. This took approximately 15 minutes.

Students then completed the PreStudy questionnaire. This took approximately 5 minutes.

Students then were asked to participate to the DescribeIT project. They were given 20 minutes to work on the image descriptions. They were then asked to complete the PreTraining Questionnaire. This took approximately 5 minutes.

Students were given 20 minutes face-to-face training about how to create good image descriptions for blind and partially sighted people by an expert in the field, Professor Helen Petrie.

Students were again asked to participate in the DescribeIT project for a further 20 minutes.

Students then completed the PostTraining Questionnaire. This took approximately 5 minutes.

Students were then debriefed about the purpose of the study and its design (as understanding study design was part of their module) and thanked for their participation.

9.2.8. Data Analysis

Many of the variables were not normally distributed. Hence, non-parametric tests were used with all the data.

9.3 Results

9.3.1 Ratings of difficulty of the image description task

Students rated how difficult the image description task would be or was at each phase of the study: before doing the task (PreStudy), after doing the task with written guidelines and an example (PreTraining) and after doing the task after verbal training (PostTraining). On each occasion, the rating was on a 7-point Likert item (1= Extremely Difficult, 7= Extremely Easy). The mean ratings and standard deviations of these three ratings are summarized in Table 9.2. To investigate if students' ratings of the difficulty of the image description task were significantly above the midpoint of the rating scale, a One-Sample Wilcoxon Signed Rank test was carried out. This showed that students' ratings in the three occasions (PreStudy, PreTraining, and PostTraining) were significantly higher than the neutral midpoint (See Table 9.2).

Table 9.2 Mean ratings (standard deviations), median and summary of significance tests of difficulty of the image description task at the three phases of the study

	Mean (SD)	**Median**	**W**	***p***
PreStudy	4.4 (1.2)	5.0	3.6	0.000
PreTraining	4.3 (1.2)	4.0	2.7	0.008
PostTraining	4.6 (1.2)	5.0	4.9	0.000

A Friedman test was conducted to test for differences between students' ratings for the difficulty level of the image descriptions task at the three phases. There was a significant difference, $\chi^2(2) = 6.2$, $p < 0.05$. A pairwise comparison showed that there were no significant differences between the PreStudy and PreTraining ratings (Z = -1.1, N=128, p=0.291) or between the PreStudy and PostTraining rating (Z = -1.5, N=128, p = 0.143), However, there was a significant difference between the PreTraining and PostTraining ratings (Z = -2.4, N=128, p <0.05), with the PostTraining ratings being significantly higher.

9.3.2 Number and quality of image descriptions

9.3.2.1 Number and quality of image descriptions produced in the PreTraining Phase

In the PreTraining Phase, overall a mean of 8.7 (SD=3.9) image descriptions were produced per student in all the conditions, including: a mean of 10.1 (SD=3.6) image descriptions were produced per student in the NoQC condition, 8.1 (SD=3.4) per student in the PeerQC condition, and 7.8 (SD=4.3) per student in the ExpertQC condition. A Kruskal-Wallis test showed that there was a statistically significant difference in the number of images produced per student in each condition, $\chi2(2) = 9.7$, p<0.05, with a mean rank of 76.2 in the NoQC condition, 58.0 in the PeerQC condition, and 53.5 in the ExpertQC condition.

The quality of the descriptions produced in the PreTraining Phase was investigated. The mean and median number of words per description in the three QC conditions are shown in Table 9.3. A Kruskal-Wallis test showed that there was a statistically significant difference in the number of words per description in each condition, $\chi2(2)$

= 19.4, p<0.001, with a mean rank of 500.3 in the NoQC condition, 588.6 in the PeerQC condition, and 582.2 in the ExpertQC condition.

Given the large variation in the lengths of the descriptions (as evidenced by the standard deviations, see Table 9.3), it was decided to divide the descriptions into word counts of different ranges (see Table 9.4) and investigate whether there was a difference in the distribution of descriptions of different lengths between the three QC conditions. A chi-square test showed there was a significant difference between the length categories of descriptions between the three QC conditions: $\chi 2$ (6) = 23.5, p<0.000.

Table 9.3 Mean (standard deviation) and median number of words per image description in the three QC Conditions in the PreTraining Phase

Condition	Mean (SD)	Median
NoQC	26.1 (21.4)	20.0
PeerQC	33.7 (28.7)	27.0
ExpertQC	31.7 (23.7)	26.0

Table 9.4 Number (percentage) of image descriptions of different lengths in the three QC Conditions in the PreTraining Phase

Word Length	NoQC	PeerQC	ExpertQC	All Conditions
1 to 25 words	287 (61)	152 (49)	154 (48)	593 (54)
26 to 50 words	116 (25)	84 (27)	103 (32)	303 (28)
51 to 75 words	51 (11)	48 (16)	44 (14)	143 (13)
76 to 100+ words	15 (3)	25 (8)	18 (6)	58 (5)
Total	469 (100.0)	309 (100.0)	319 (100.0)	1097

Further analysis of the quality of 96 descriptions (of the non-trivial descriptions) was conducted (see section 7.2.1.5 for details of calculation of these measures). The mean and median ratings and standard deviations of the accuracy, error rate, and efficiency of the image descriptions are summarized in Table 9.5. Kruskal Wallis tests showed no significant differences in these variables between the three QC conditions. Therefore, further analysis of the accuracy, error score, and efficiency of the image descriptions produced in this Phase will combine data from the three QC conditions together.

Experts initially decided whether an image description was suitable as a short or long description (see section 7.2.1.5 for further detail). The mean efficiency score of the short descriptions was 19.1 (SD =10.1), and for the long descriptions it was 15.2 (SD =7.1). A Mann-Whitney U test was used to compare the experts' decision with the efficiency scores of the descriptions. It showed that the efficiency scores of the descriptions the experts decided were appropriate as short descriptions were significantly higher than for those descriptions which the experts decided were appropriate as long descriptions ($U = 721$, $p < 0.05$).

Table 9.5 Mean (standard deviation) and median accuracy, error score, and efficiency of sample of image descriptions for the three QC conditions in the PreTraining Phase

Quality Measure	**NoQC**		**PeerQC**		**ExpertQC**		**Overall**	
	Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
Accuracy	59.3 (24.9)	59.3	61.3 (24.0)	62.5	51.0 (29.0)	53.0	57.1 (26.1)	57.1
Error score	0.3 (1.5)	0.0	0.0 (0.0)	0.0	0.7 (3.1)	0.0	0.3 (2.0)	0.0
Efficiency	20.2 (9.5)	19.3	17.9 (8.7)	18.1	15.0 (9.2)	14.0	18.0 (9.3)	17.0

Experts also rated the quality of the 10% sample of descriptions on a 5-point Likert item (1=Very poor, 5=Excellent). Table 9.6 shows the mean and median quality ratings for the three QC conditions. A Kruskal Wallis test showed no significant difference between the experts' ratings in each condition.

Table 9.6 Mean (standard deviation) and median expert ratings of image description quality for the three QC conditions in the PreTraining Phase

NoQC		**PeerQC**		**ExpertQC**		**Overall**	
Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
3.5 (0.9)	3.8	3.5 (0.9)	3.5	3.0 (1.3)	3.5	3.3 (1.1)	3.5

Most relevant measure of quality was the experts' ratings and the accuracy score. Hence, Spearman correlations were conducted to investigate if there were relationships between expert ratings and accuracy scores of the descriptions. There was a significant positive correlation between the accuracy and expert ratings ($r_s(96) = 0.57$, $p < 0.000$). There was no correlation between Error score and experts' ratings.

Table 9.7 presents some examples of experts' comments about the quality of the image descriptions with the associated rating.

Table 9.7 Examples of experts' comments about their ratings of image descriptions for different levels of quality rating in the PreTraining Phase

Rating	Expert comments
1	*The description does not explain the image at all, just the slide's background and text content.* (Expert 7) *Too much information about the slide. Only the last part of the first line of this description is relevant and provides details about the mice, their location and dialogue.* (Expert 6)
2	*Again, this description does not only describe the imabge but the whole slide, and the bit that does refer to the image might make quite a good short description.* (Expert 5) *This includes the text outside of the image - perhaps suitable as a short description if you removed the text bits* (Expert 2)
3	*It's a clear description of the drawing, but does not indicate the implied meaning from the drawing - i.e. that the cylinder is representative of a glass which is both half full and half empty.* (Expert 7) *There's no indication that the image is, essentially, decorative and does not actually convey any real content. The description is fairly detailed and accurate but doesn't convey what it actually represents.* (Expert 4)
4	*This describer has captured the essence of the image that complements the text. It would have a 5 if he had described the books better.* (Expert 1) *Just enough detail to tell the story and make the point.* (Expert 2)
5	*The description highlights the key features of the image, including getting across the emotion portrayed by the woman.* (Expert 7) *This is a detailed description of the picture providing many details about the person, the objects around him and the environment.* (Expert 6)

9.3.2.2 Number and quality of image descriptions produced in the PostTraining Phase

In the PostTraining Phase overall a mean of 7.7 (SD=3.9) image descriptions were produced per student in all the conditions, including: a mean of 8.1 (SD=4.2) image descriptions were produced per student in the NoQC condition, 7.8 (SD=3.9) per student in the PeerQC condition, and 7.2 (SD=3.8) per student in the ExpertQC condition. A Kruskal Wallis test showed no significant differences between the number of images per student produced in the three QC conditions.

The quality of the descriptions produced in the PostTraining Phase was investigated. The mean and median number of words per description in the three QC conditions are shown in Table 9.8. A Kruskal-Wallis test showed that there was a statistically significant difference in the number of words per description in the CQ conditions, $\chi2(2) = 7.7$, $p<0.05$, with a mean rank of 430.6 in the NoQC condition, 483.3 in the PeerQC condition, and 479.3 in the ExpertQC condition.

Table 9.8 Mean (standard deviation) and median number of words per image description in the three QC Conditions in the PostTraining Phase

Condition	Mean (SD)	Median
NoQC	28.4 (17.9)	26.0
PeerQC	32.9 (20.7)	30.0
ExpertQC	33.4 (23.0)	30.0

Although the variation in the lengths of the descriptions were not as large as in the PreTraining Phase (compared the standard deviations in Table 9.3 with the standard deviations in Table 9.8), it was decided to divide the descriptions into word counts of different ranges (see Table 9.9) and investigate whether there was a difference in the distribution of descriptions of different lengths between the three QC conditions. A chi-square test showed there was a significant difference between the length categories of descriptions between the three QC conditions: $\chi2\ (6) = 16.31$, $p<0.05$. Inspection of the percentages shows that the percentage of long descriptions (over 75 words) was substantially higher in the Peer and Expert QC conditions.

Table 9.9 Number (percentage) of image descriptions of different lengths in the three QC Conditions in the PostTraining Phase

Word Length	NoQC	PeerQC	ExpertQC	Total
1 to 25 words	163 (49)	138 (46)	116 (41)	417 (46)
26 to 50 words	125 (38)	103 (34)	110 (39)	338 (37)
51 to 75 words	40 (12)	47 (16)	40 (14)	127 (14)
76 to 100+ words	3 (1)	13 (4)	17 (6)	33 (4)
Total	331 (100)	301 (100)	283 (100)	915 (100)

Further analysis of the quality of 96 descriptions (of the non-trivial descriptions) was conducted (see section 7.2.1.5 for details of calculation of these measures). The mean and median ratings of the accuracy, error score, and efficiency of the image descriptions are summarized in Table 9.10. Kruskal Wallis tests showed no significant differences in these variables between the three QC conditions. Therefore, further analysis of the accuracy, error score, and efficiency of the image descriptions produced in this Phase will combine data from the three QC conditions together.

Table 9.10 Mean (standard deviation) and median accuracy, error rate, and efficiency of sample of image descriptions for the three QC conditions in the PostTraining Phase

Quality Measure	NoQC		PeerQC		ExpertQC		Overall	
	Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
Accuracy	72.1 (26.0)	79.0	64.9 (25.3)	62.0	70.4 (25.1)	73.3	69.1 (25.4)	71.4
Error score	1.0 (2.8)	0.0	0.0 (0.0)	0.0	0.0 (0.0)	0.0	0.2 (1.2)	0.0
Efficiency	19.0 (8.1)	18.0	19.3 (8.3)	18.0	19.0 (9.0)	20.0	19.0 (8.3)	18.0

Experts initially decided whether an image description was suitable as a short or long description (see section 7.2.1.5 for further detail). The mean efficiency score of the short descriptions was 21.8 (SD =7.1), and for the long descriptions it was 19.0 (SD=6.6). A Mann-Whitney U test was used to compare the experts' decision with the efficiency scores of the descriptions. It showed that the efficiency scores of the descriptions the experts decided were appropriate as short descriptions were significantly higher than for those descriptions which the experts decided were appropriate as long descriptions (U = 707, p <0.05).

Experts also rated the quality of the 10% sample of descriptions on a 5-point Likert item (1=Very poor, 5=Excellent). Table 9.11 shows the mean and median quality ratings for the three QC conditions. A Kruskal Wallis test showed no significant difference between the experts' ratings in each condition.

Spearman correlations were conducted to investigate if there were relationships between expert ratings and accuracy scores of the descriptions. There was a significant positive correlation between the accuracy and expert ratings ($r_S(96) = 0.27$, $p < 0.01$). No significant correlation was found between expert ratings and error rate score.

Table 9.11 Mean (standard deviation) and median expert ratings of image description quality for the three QC conditions in the PostTraining Phase

NoQC		PeerQC		ExpertQC		Overall	
Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
3.8 (0.7)	4.0	3.8 (0.8)	4.0	3.8 (0.7)	4.0	3.8 (0.7)	4.0

9.3.2.3 Comparison of number and quality of image descriptions between the PreTraining and PostTraining Phases

To investigate whether the training had an effect on the number of image descriptions produced, related-samples Wilcoxon signed rank tests were conducted. The results of these tests showed a significant decrease of number of images produced in NoQC condition ($Z=-3.0$, $p<0.01$). However, there was no differences in the number of image descriptions produced in the PreTraining and PostTraining Phases in both PeerQC and ExpertQC conditions.

Another series of related-samples Wilcoxon signed rank tests were conducted to investigate whether the training had an effect on the number of words per description in the QC conditions. The results showed no significant differences.

Lastly, to investigate whether the training had an effect on the quality of the image descriptions produced, related-samples Wilcoxon signed rank tests were conducted on the accuracy scores, error score, efficiency scores, and experts' ratings of the quality of the image descriptions produced in the PreTraining and PostTraining Phases. The results of these tests are summarized in Table 9.12. Accuracy and efficiency scores and experts' ratings were all significantly higher in the PostTraining Phase in comparison to the PreTraining Phase.

Table 9.12 Results of related-samples Wilcoxon signed rank tests for accuracy scores, error rate, efficiency scores, and experts' ratings of quality for descriptions produced in the PreTraining and PostTraining sessions

Measures	**W**	***p***
Accuracy	3.2	0.001
Efficiency	2.6	0.010
Error score	-0.9	n.s.
Expert ratings	2.8	0.005

9.3.3 Students' motivation to participate in the image description task

To investigate students' motivation to participate in the image description task, they completed the Situational Motivation Scale (SIMS) after undertaking the task in the PreTraining Phase and again after undertaking the task in the PostTraining Phase. The mean (standard deviation) and media scores on the four sub-scales of the SIMS for the three QC conditions are summarized in Table 9.13. for the PostTraining Phase and Table 9.14 for the PostTraining Phase.

A series of Kruskal Wallis tests showed no significant difference between the SIMS sub-scales in each condition in both the PreTraining and PostTraining Phases.

Table 9.13 Mean (standard deviation) on SIMS subscales in the three QC conditions in the PreTraining Phase

SIMS subscale	NoQC		PeerQC		ExpertQC	
	Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
Intrinsic Motivation	3.2 (1.5)	3.0	3.5 (1.3)	3.5	3.1 (1.5)	3.0
Identified Regulation	2.6 (1.3)	2.0	2.6 (1.4)	2.0	2.7 (1.5)	2.0
External Regulation	5.0 (1.4)	5.0	5.0 (1.6)	5.0	4.7 (2.0)	5.0
Amotivation	2.6 (1.4)	2.0	3.0 (1.4)	3.0	2.8 (1.4)	3.0

A series of two sample related tests were conducted to investigate whether the training had an effect on students' motivation. The results showed a significant effect on the Intrinsic Motivation sub-scale (Z= -3.4, $p<0.01$), where students intrinsic motivation significantly decreased.

To investigate whether there was a relationship between number of images produced and situational motivation, Spearman correlations were calculated between the number of images produced in both the PreTraining and PostTraining Phases and the four sub-scales of the SIMS. None of these correlations were significant in the PreTraining phase. However, in the PostTraining Phase a significant positive correlation was found between the number of images produced per student and the SIMS sub-scale Extrinsic Motivation ($r_s(90) = 0.23, p<0.05$).

Table 9.14 Mean (standard deviation) and median on SIMS subscales in the three QC conditions in the PostTraining Phase

SIMS subscale	NoQC		PeerQC		ExpertQC	
	Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
Intrinsic Motivation	2.8 (1.4)	2.5	3.3 (1.3)	3.0	2.9 (1.5)	2.0
Identified Regulation	2.5 (1.4)	2.0	2.8 (1.5)	3.0	2.7 (1.6)	2.0
External Regulation	4.5 (1.5)	5.0	4.9 (1.6)	5.0	4.6 (2.1)	5.0
Amotivation	2.5 (1.4)	2.0	3.1 (1.3)	3.0	3.0 (1.7)	3.0

Another Spearman correlations tests were conducted between the number of words per descriptions in both the PreTraining and PostTraining Phases and the four sub-scales of the SIMS. In the PreTraining phase this showed a significant positive correlation between the Intrinsic Motivation sub-scale and the number of words ($r_s(99)=0.24, p<0.05$), and a significant negative correlation between the Amotivation sub-scale and the number of words ($r_s(99) = -0.29, p<0.01$). Similar results were found in the PostTraining Phase, where a significant positive correlation between the Intrinsic Motivation sub-scale and the number of words ($r_s(90)=0.26, p<0.05$) was found, and a significant negative correlation between the Amotivation sub-scale and the number of words ($r_s(90) = -0.25, p<0.05$).

9.3.4 Effect of attitudes to people with disabilities and altruism on students' motivation to participate in the image description task

To investigate whether students attitude to people with disabilities and their sense of altruism had an effect on their motivation to participate in the image description task, they completed the Interaction with Disabled People Scale (IDP) and the Self-Altruism Scale (SRA) at the end of the PostTraining Phase. The mean (standard deviations) and median scores for both scales for students in the three QC conditions are summarized in Table 9.15.

Table 9.15 Mean (standard deviation) and median on IDP and SRA subscales in the three QC conditions

Measures	NoQC		PeerQC		ExpertQC	
	Mean (SD)	**Median**	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
IDP	69.3 (10.7)	71.0	67.7 (13.4)	66.0	69.7 (9.4)	70.0
IDP Sub-scales						
Sympathy	23.1 (4.1)	24.0	21.5 (5.3)	22.0	22.3 (3.6)	23.0
Discomfort	12.0 (4.4)	11.5	10.0 (5.2)	10.0	12.7 (4.9)	13.0
SRA	38.7 (7.4)	39.0	37.4 (8.2)	38.0	38.4 (8.6)	38.0

The overall mean IDP score was 68.9 (SD= 11.2). A comparison with Brown et al. (2009) reported results on the IDP for British undergraduate students was made. A one-sample Wilcoxon Signed Rank test found that the scores of the students in this study was significantly higher than the median of IDP scores provided by Brown et al (2009) (W=5.0, p<0.001).

The overall mean SRA score was 38.1 (SD= 8.1). A one-sample Wilcoxon Signed Rank test found that the scores of MTurk workers was significantly higher than the median of SRA scores provided by Rushton et al. (1981) (W=4.1, p<0.001).

To investigate whether there was a relationship between the number of image descriptions produced (in both the PreTraining and PostTraining Phases) and the scores on the IDP and SRA, a series of Spearman correlation tests were conducted. None of these correlations were significant.

Another series of Spearman correlation tests were conducted to investigate whether there was a relationship between the number of words (in both the PreTraining and PostTraining Phases) and the scores on the IDP and SRA. Results are summarized in Table 9.16. In the PreTraining phase none of these correlations were significant. However, in the PostTraining phase a positive significant correlation was found in the IDP score ($r_s(90) = 0.23, p<0.05$) and discomfort sub-scale ($r_s(90) = 0.27, p<0.05$). No significant correlation found between SRA and Sympathy sub-scale.

Table 9.16 Spearman correlations between IDP and SRA scales and the number of words per descriptions in PreTraining and PostTraining Phases

Measures	**PreTaining**	**PostTraining**
IDP Scale	n.s.	0.030
Sympathy	n.s.	n.s.
Discomfort	n.s.	0.011
SRA Scale	n.s.	n.s.

9.4 Discussion and Conclusions

This study investigated the effect of a face-to-face training session on students' motivation to participate in the image description task and particularly on the quantity and quality of the images produced. It also investigated the effects of different quality control messages in the instructions for the image description task. Finally, it investigated students' perceptions of the difficulty of the image description task and whether this was altered by the training. The study confirmed some of the hypotheses but failed to support others. The Study summary presented in Table 10.17 end of this chapter.

- *Does a face-to-face training session decrease students' perception of the difficulty of the image description task?*

The results of this study found that even before describing any images students perceived the image description task as easy task to do (they rated it significantly above the midpoint on the difficulty item, so towards the easy pole of the item). However, after the training and then again trying the image description task in the PostTraining phase, students thought that the image description task was significantly easier than the PreTraining phase, but not significantly different from their initial difficulty ratings.

Thus, providing the training did significantly decrease the perception of the difficulty of the task, in comparison to when students undertook the task without training. Hence, providing training (this could be in the form of a video rather than face-to-face training) could improve the participation rate in this kind of crowdsourcing project.

Building on these results, the low participation rate and total lack of participation in Studies 2, 3, and 4 cannot be explained by the difficulty level of the image description task. Students in the current study found the task easy before describing any images and again in the PreTraning Phase. Accordingly, students in Studies 2, 3, and 4, might have thought they would participate in the project (as the results of Phase 1 suggested) because they thought providing image descriptions is an easy task, but their failure to provide descriptions does not seem to be because they found the description task difficult once they tried it.

- *Do instructions about quality control (i.e. that a peer or an expert will check the quality of the descriptions) increase the quality of the descriptions produced?*

Both in the PreTraining and PostTraining phases there was a statistically significant difference in the number of words per description between the three QC conditions. There were longer descriptions in the PeerQC and ExpertQC conditions in comparison to the NoQC conditions. There was also a significant difference in the length categories of descriptions between the three QC conditions, with the percentage of long descriptions (over 75 words) being substantially higher in the Peer and Expert QC conditions.

However, further analysis of the 10% sample of the descriptions collected in each phase (PreTraining and PostTraining) showed that there were no significant differences between the experts' ratings of quality in the three conditions. In addition, there were no differences in accuracy, error rate, and efficiency between the three QC conditions.

This means that although the PeerQC and ExpertQC conditions produced longer descriptions compared to the NoQC condition, the overall quality was not different across the three instructions conditions either before or after training. In fact, the descriptions that experts decided were suitable for short descriptions had significantly higher efficiency in both PreTraining and PostTraining phases.

- *Does a face-to-face training session improve students' motivation to participate in the image description task and improve the quantity and quality of the image descriptions produced?*

The training had a significant effect on students' motivation as the Intrinsic Motivation sub-scale significantly decreased. However, this was not accompanied with increase in Extrinsic Motivation. Some researchers argue that intrinsic and extrinsic motivation are the extreme point of the same continuum (e.g. Harter 1981) so this result would be surprising. However, other researchers argue that intrinsic and extrinsic are independent of each other (Deci and Ryan 1985, Vallerand 1977), so this result is possible.

The number of images produced per student significantly decreased in the PostTraining phase, and although the average number of words per description did not significantly change, the accuracy and efficiency of the descriptions did significantly increase. This means that although students described less images in the PostTraining phase the quality of these descriptions was significantly higher.

Similar to the results of Studies 5 (in the InMot condition), 6, and 7 there was no correlation between number of images produced and situational motivation of the students in the PreTraining phase. However, there was a significant positive correlation between the number of images produced per student and scores on the SIMS Extrinsic Motivation sub-scale in the PostTraining Phase. This result is interesting because there was no significant change between the PreTraining and PostTraining Phases in the SIMS Extrinsic Motivation sub-scale.

Furthermore, there was a significant positive correlation between the Intrinsic Motivation sub-scale and the number of words per description, and a significant negative correlation between the Amotivation sub-scale and the number of words per description in both the PreTraining and PostTraining Phases. This means that students who were intrinsically motivated in both the PreTraining and PostTraining Phases produced longer descriptions. In contrast students who were not motivated in any way produced shorter descriptions, probably because they did not value the image description task or the outcome of their participation.

- *Do levels of altruism correlate with the quality and quantity of the descriptions produced?*
- *Do attitudes to people with disabilities correlate with the quality and quantity of the descriptions produced?*

The level of participation was measured by the number of images described by each student and the quality of the descriptions produced. It was hypothesized that there would be a positive relationship between both the number of images described and the number of words per description with students' sense of altruism, and attitudes towards disabled people (discomfort and sympathy) in both the PreTraining and PostTraining Phases. The study showed that there was no relationship between students' general sense of altruism and either the quality or quantity of the image descriptions they produced in both the PreTraining and PostTraining phases. Although the relationship between students' sense of altruism and quality of the descriptions was not investigated in the previous studies (Studies 5 to 9), the relationship between students' sense of altruism and the number of images described per student was investigated and yielded similar results, in that in general there was no significant correlation between self-reported altruism and the number of image descriptions produced. The exception was for the ExMot condition in Study 5 in which there was a significant positive correlation.

There was no significant correlation between attitudes to disabled people (as measured by the IDP) and the quantity of the image descriptions produced in either the PreTraining and PostTraining Phases. In general, the results of the PreTraining Phase are similar to those found in Studies 5 to 9. The exceptions are in Study 6 in which there was a negative correlation with the Discomfort sub-scale of the IDP, and in Study 5 (in the ExMot condition) in which there was a positive correlation with the Sympathy sub-scale. In addition, there was no significant correlation between attitudes to disabled people and the quality of the image descriptions produced in the PreTraining Phase. Interestingly, in the PostTraining phase there was a positive significant correlation between the number of words and attitudes to disabled people. This means that the more students felt discomfort towards disabled people, the longer their descriptions were after training.

One limitation of the study is the setting in which the study was run, a practical class for undergraduate students. For example, the significant decrease in the intrinsic motivation in the task in the PostTraining Phase may have been because students were becoming irritated with having to do the same task again, rather than the effect of the training.

In conclusion, this study showed that a face-to-face training session to teach students how to provide good image descriptions to blind and partially sighted students and to emphasise the importance of this task resulted in making the task easier for students than only presenting them with a set of guidelines. In addition, it led to significant improvement in the quality of the descriptions produced in comparison to presenting students only with the set of guidelines.

The Quality Control information in the instructions in the PeerQC and ExpertQC conditions resulted in producing longer descriptions in compare to NoQC, but had no effect on the accuracy and efficiency of the descriptions across the three conditions.

Table 9.17 Study 10 summary

- There was a significant difference between the students' ratings for the difficulty level of the image descriptions task at PreTraining and PostTraining ratings with the PostTraining ratings being significantly higher.
- The PreTraining Phase:
  1. There was a statistically significant difference in the number of images produced per student in each condition.
  2. There was a statistically significant difference in the number of words per description in each condition.
  3. There was a significant difference between the length categories of descriptions between the three QC conditions.
  4. There were no significant differences in accuracy, error score, and efficiency between the three QC conditions.
  5. The efficiency scores of the descriptions the short descriptions were significantly higher than long descriptions.
  6. There was no significant difference between the experts' ratings in three conditions.
  7. There was a significant positive correlation between the accuracy and expert ratings.
- PostTraining Phase:
  1. There were no significant differences between the number of images per student produced in the three QC conditions.
  2. There was a statistically significant difference in the number of words per description in the CQ conditions.
  3. There was a significant difference between the length categories of descriptions between the three QC conditions.
  4. There were no significant differences in accuracy, error score, and efficiency between the three QC conditions.
  5. The efficiency scores of the descriptions the short descriptions were significantly higher than long descriptions.
  6. There was no significant difference between the experts' ratings in three conditions.
  7. There was a significant positive correlation between the accuracy and expert ratings.
- Comparison between the PreTraining and PostTraining Phases:

1. There was a significant decrease of number of images produced in NoQC condition.

2. There was no significant difference between the number of words per description in the QC conditions.
3. Accuracy and efficiency scores and experts' ratings were all significantly higher in the PostTraining.

- Students' motivation:
  1. There was no significant difference between the SIMS sub-scales in each condition in both the PreTraining and PostTraining Phases.
  2. The training had a significant effect on students' motivation where the Intrinsic Motivation sub-scale significantly decreased.
  3. No correlation between number of images produced and situational motivation in the PreTraining phase.
  4. There was a significant positive correlation between the number of images produced per student and the SIMS sub-scale Extrinsic Motivation in the PostTraining Phase.
  5. There was a significant positive correlation between the Intrinsic Motivation sub-scale and the number of words, and a significant negative correlation between the Amotivation sub-scale and the number of words both in the PreTraining and PostTraining Phases.
- Attitude toward disabled and sense of altruism:
  1. There were no significant correlations between the number of image descriptions produced (in both the PreTraining and PostTraining Phases) and the scores on the IDP and SRA.
  2. There was no significant correlation between the number of words and the scores on the IDP and SRA in the PreTraining phase.
  3. There was a positive significant correlation between the number of words and the scores on the IDP, but no significant correlation was found with the SRA score in the PostTraining Phase.

Chapter 10

Overall Discussion and Conclusions

Crowdsourcing has now been used in many areas, both in the private sector and for socially responsible purposes. One socially responsible area is to support people with disabilities. However, there is little exploration of what motivates people to participate in this latter type of crowdsourcing project. The initiatives and projects reported in Chapter 2 (section 2.3) have shown that crowdsourcing has the potential to be a useful means of undertaking socially responsible activities to support people with disabilities. However, the research on crowdsourcing projects to support people with disabilities has not investigated what motivates people to participate in these projects. This programme of research aimed to investigate what would motivate university students to participate in a socially responsible crowdsourcing project, the DescribeIT, to support blind and partially sighted students by describing images in digital learning resources. More specifically, the programme of research aimed to investigate how different motivational factors can affect the participation of sighted students in the DescribeIT project and how they can affect the quantity and quality of the descriptions produced.

In this final Chapter, the key findings of this programme of research are discussed first by revisiting the research questions in light of Self-determination Theory and Fogg's Behavioural Change Model. In addition, the contributions and the limitations of the research is presented.

10.1 Research Questions and Research Findings

The key research question in this programme of research was "What are the motivators for students to participate in socially responsible crowdsourcing projects?", the hypothesis was that different motivational factors have different effects on students' participation rate and the quantity and quality of the image descriptions produced. Initially students' self-reports of motivation were investigated. Some of the motivational factors found in the self-report motivation studies were incorporated into the DescribeIT project. A student's decision to contribute to the DescribeIT project meant that the student is motivated to participate to the project to support disabled students. The Situational Motivation Scale (SIMS) was employed to measure students' motivation while participating in the DescribeIT project. The SIMS measures Intrinsic

Motivation, Identified Regulation, External Regulation, and Amotivation. The quantity and quality of the image descriptions produced were also measured to assess students' level of participation in the image description task. It was also hypothesised that students' altruistic attitudes is an important driver of participation since the DescribeIT project was designed to be a volunteer-based project. Furthermore, it was hypothesised that students' attitudes towards disabled people would influence their participation to the DescribeIT project.

In the following sections I will highlight the main findings of this programme of research. The aim is not to present in detail the findings discussed at the end of each chapter, rather it is to highlight the overall findings.

1. **Students' Likelihood of Participating in the DescribeIT Project**

Overall, in the self-report phases of the studies students showed a positive interest in participating in the DescribeIT project to describe images in electronic learning resources for blind and partially sighted students. Students rated the likelihood of their own participation significantly higher than that of other students on their course. This could be due to what social psychologists call the "fundamental attribution error" (Ross and Nisbett, 2011) which proposes that people interpret their own behaviour very much in terms of the specific situation, but interpret the behaviour of others in terms of persistent personality traits. Thus, when asked to predict their own participation, students think of the specific situation of helping other students who are at a disadvantage, but when asked to predict the behaviour of other students, they think of the general helpfulness of other students, not in the context of helping disabled students

2. **Students' Self-Reports of Motivation to Participate in the DescribeIT Project**

Studies 1 to 6 presented a first insight of what students thought would motivate them to participate in a crowdsourcing project to support disabled students. The studies revealed that altruism and monetary rewards were the leading self-reported motivational factors for UK students, whereas, altruism and wanting to support research were the leading self-reported motivational factors for Arab students. Students suggested other intrinsic and extrinsic motivational factors that would motivate them to participate to the DescribeIT project, such as being interested in the accessibility area, the time expected from them to participate in the project being not too long, the ease of providing the descriptions, seeing how their participation would be helpful to blind or partially sighted students, participation increasing their understanding of the learning resources and knowing a blind or partially sighted person. These findings reinforce the notion that each crowdsourcing project has its own particular relevant motivations, although in line with previous research (See Chapter 2, section 2.5), these studies also showed that there are numerous self-reported motivations in common.

A triangulation between what students said would motivate them to participate in the DescribeIT project (an open-ended question) and their ratings of the motivational factors showed the influence of social desirability bias in students' answers. For example, overall students' answers to the open-ended question showed that monetary rewards is the most mentioned motivational factor, followed by altruism (except for Studies 5 and 6). However, ratings of the motivational factors showed that altruism had the highest mean, not monetary rewards. In addition, a triangulation between what students said would motivate them and what would motivate other students showed the influence of social desirability bias. Students generally emphasized the importance of altruism to themselves in comparison to other students, and emphasized the importance of monetary rewards to other students in comparison to themselves.

3. Intrinsic Motivation vs. Extrinsic Motivation – Students' Actual Behaviour

In line with the results of the triangulation made in the self-report motivation phases of the studies, the very low participation rates in Studies 2 to 4, when students were asked to actually provide descriptions, provide empirical evidence of the magnitude of the social desirability bias in reporting motivations to participate in socially responsible crowdsourcing projects. By considering two main findings: (1) Students interest in participating in the DescribeIT project (as reported in Phase 1) and (2) their self-report of the importance of altruism (as reported in Phase 1), it is clear that very few students want to say that they are unwilling to help disabled students, even in an anonymous questionnaire. In addition, perhaps students believe that it is morally required from them to support disabled students, so they find it difficult to admit even to themselves that they are not motivated by altruism to support disabled students (i.e. they would like to think they are motivated by altruism).

Table 10.1 summarize the participation in the DescribeIT project in Studies 2 to 9. Whilst intrinsic and non-financial extrinsic motivational factors did not motivate UK students to participate in the DescribeIT project (Studies 2 to 4), the extrinsic motivational factor of monetary reward did motivate UK students to participate (Study 6). Interestingly, the same groups of students who participated in Study 6 were invited to participate in Studies 2 to 4 but did not participate (when non-financial motivation were highlighted). By revisiting the result on monetary rewards in the self-report of motivation phase of students of Study 6, one would conclude that students' actual behaviour when monetary was offered is in line with their ratings on the monetary reward factor (significantly above the midpoint of the scale).

The interviews with a small number of the students in Study 3 showed a range of sensible reasons for students' lack of participation, including pressure of studying, simple forgetfulness, being busy with other research and finding the task daunting. However, in addition to the above reasons mentioned by students to explain their lack of participation, perhaps it is important to consider two findings from the self-reports phases of the initial studies in light of Fogg's Behavioural Change Model, to understand students' behaviour when they were asked to participate to the DescribeIT project. The first finding is that students in Studies 1 to 5 thought they would not be motivated at all by passing the time. The second finding is that, students repeatedly mentioned that they would participate to the project if their participation does not require too much of their time. According to Fogg's Behavioural Change Model, not having the time to perform a particular behaviour decreases a person's ability to do the behaviour. In the context of Studies 2 to 4, students wanted to participate in the DescribeIT project (as the self-reports showed), but not having the time due to being busy with other research, or their own work (lectures, assignments and exams) as reported in Study 3 might have had an effect in reducing their ability to participate. On another point, according to the model even if students perceived the description task as an easy task to do (as suggested by Study 10), not having the time to do the task means that the task is no longer easy to do.

In Study 5, Arab students were highly motivated to participate in the DescribeIT project, and intrinsic motivations dominated over extrinsic motivations. The high participation rate in Study 5 in comparison to Studies 2 to 4 and 6 suggests a cross-cultural difference between what would motivate UK students and Arab students to participate in a crowdsourcing project to support blind and partially sighted students. Further research is needed to confirm this. If these differences are confirmed, then crowdsourcers should take in consideration the diversity of what would motivate people from different cultures when recruiting a crowd for a particular task.

Lastly, it is worth noting that the highest participation rate found in the DescribeIT project with non-financial motivational factors was in Study 2 (11%) and Study 5 (30%). In Study 2 all the participants were postgraduate students (100%) and in Study 5, over 60% of the participants were postgraduate students. In addition, in Study 8 with the TagIT project only postgraduate students participated, although undergraduate and postgraduate students were invited to take part in the study (34% of the postgraduate students invited participated, compared to an overall participation rate of 24%). This result suggests that postgraduate students are motivated more than undergraduate students to participate for non-monetary reasons. However, it should be noted that the majority of participants in Study 2 were international students (UK students - 23%, international students - 77%), a very diverse group in terms of culture. In Study 8 only 11 students completed the demographic section, and students were from UK, India, and China. Thus, there could be a cultural factor, which was not controlled for, that had an effect on students' participation rate, which was also suggested by the findings of Study 5.

4. **Quantity and Quality of Image Descriptions by Students vs Committed Crowd Members**

The total number of images produced in each study and the mean number of images described per participant is summarized in Table 10.1. A Kruskal-Wallis test showed that there was a significant difference in the number of image descriptions produced per participant between the studies 5 (InMot condition), 5 (ExMot condition),6,7, and 9, $\chi2(4) = 84.8$, $p < 0.001$, with a mean rank of 221.4 for Study 6, 192.0 for Study 5(InMot condition), 178.0 for Study 5(ExMot condition), 106.4 for Study 7 and 100.6 for Study 9. Studies 2 to 4 were not included in this comparison due to the very low number of participants in comparison to studies 5 to 9. This means that when UK students were paid to do the description task they produced a higher number of images per student in comparison to when Arab students were asked to describe images using either an intrinsic or an extrinsic motivation. Interestingly, although MTurk workers produced the highest number of images overall, the number of images produced per worker was only the fourth highest after those of Study 6 and Study 5 (InMot and ExMot conditions). Crowdcrafter members produced the lowest number of images per participant and had the lowest total number of images.

Table 10.1 Participation measures in Studies 2 to 9

Measures	Study 2		Study 3		Study 4	Study 5		Study 6	Study 7	Study 9
	InMot	ExMot	InMot	ExMot	InMot	InMot	ExMot	ExMot	ExMot	ExMot
Number of Participants	5	3	0	4	2	44	45	26	164	32
Participation Rate	14%	8%	0%	3%	0.8%	29%	30%	24.5%	-	-
Number of images	83	3	0	17	11	238	346	385	851	102
Average number of images per participant (SD)	7.8 (10.0)	1.0 (0.0)	0.0 (0.0)	4.3 (5.3)	5.5 (0.7)	9.4 (3.9)	7.9 (3.8)	17.5 (9.1)	5.6 (8.5)	4.3 (6.9)

To investigate the quality of the image descriptions created in a variety of contexts, comparisons between samples of 10% of the image descriptions were made. As shown in Table 10.2, comparisons between the quality of a sample of the descriptions produced by UK students in Study 6, the quality of sample of the descriptions produced by MTurk workers in Study 7 and the quality of a sample of the descriptions produced by CrowdCrafting members in Study 9 were made. The Kruskal-Wallis test showed that there no significant difference in term of experts' ratings, error score and accuracy of the descriptions. However, descriptions of MTurk workers were significantly lower in efficiency than those produced by students and CrowdCrafting members, $\chi2(2) = 23.0$, $p < 0.001$, with a mean rank of 90.4 for CrowdCrafting members, 83.0 for students, and 53.0 for MTurk workers. The image descriptions produced by Arab students in Study 5 were only investigated in terms of word count and the relation between word count and number of objects mentioned in the description. The results showed that the majority of the descriptions were 50 words or less, which means that these descriptions are probably not detailed enough for blind and partially sighted students. The number of objects mentioned in these descriptions were significantly correlated with the word count of the descriptions, which mean the longer the descriptions the more details it was. This means that the quality of the descriptions in terms of accuracy, error score, and expert rating was not affected when different motivational factors were used with different groups of participants (students or crowd members). However, the efficiency level was significantly lower with the MTurk workers, perhaps workers thought the longer the descriptions is the higher quality would be.

However, it should be highlighted that in Studies 6, 7, and 9, the further analysis of the quality of the descriptions was assessed to only 10% of the descriptions produced in each study. Hence, the number of images from each study was not balanced. In addition, different measures were used in Study 5. Thus, the comparison of the quality of the image descriptions produced from these studies is not ideal, nevertheless it is helpful to establish the overall levels of quality of image descriptions obtained from students and crowd members.

Table 10.2 Image description quality measures in Studies 5 to 9

Measures	Study 5		Study 6	Study 7	Study 9
	InMot Condition	ExMot Condition	ExMot Condition	ExMot Condition	ExMot Condition
Percentage of descriptions which were 50 words or less	88%	97%	69%	77%	94%
Percentage of descriptions which were 51 words or more	12%	3%	31%	23%	6%
Quality of descriptions (10% sample)					
Accuracy	-	-	49.3%	52.9%	58.5%
Efficiency	-	-	21.0%	14.3%	21.6%
Error Score	-	-	0.0%	1.6%	0.0%

5. Differences in motivation between committed crowd members and students

These studies have shown that the motivation of crowd members (MTurk and CrowdCrafting) and university students are different. People visit crowdsourcing platform explicitly to complete tasks, hence crowd members in Studies 7 and 9 had the sense that they were voluntarily participating in the DescribeIT project, this was assessed by the SIMS sub-scale Identified Regulation. Whereas, students who were invited to participate in the DescribeIT project in Studies 5 and 6 were neutral in their feelings that they were voluntarily choosing to participate to the project. This was despite the fact that the students in studies 5 and 6 voluntarily took part in the studies. The same results were found when Crowdcrafting members and students participated in the TagIT project in Study 9 and Study 8 respectively. As measured by the Identified Regulation sub-scale of the SIMS, Crowdcrafting members in Study 9 (TagIT project) had the sense that they were voluntarily participating, whereas students in Study 8 were neutral in their feelings that they were voluntarily choosing to participate to the TagIT project.

6. Overall financial cost and time to produce a set of image descriptions for blind and partially sighted students

If the idea of crowdsourcing the creation of image descriptions for blind and partially sighted students is to become a reality, the cost in terms of both money and time to produce image descriptions needs to be carefully considered. This is particularly important in light of the fact that with the UK students, the non-monetary rewards failed to produce participation rates which would sustain the creation of image descriptions for the many digital learning resources used in institutions of higher education. As shown in Table 10.3, in Study 6 the 385 descriptions produced by UK students cost approximately GBP 80 (or 20p per description), in comparison the 851 descriptions produced by MTurk workers in Study 7 cost approximately GBP 82 (9.6p per description – as a 20% administration fees must be paid to MTurk in addition to the 8p per description paid to the MTurkers). In Study 5 Arab students produced a total of 584 image descriptions and in Study 9 the crowd members of CrowdCrafting platform produced a total of 102 image descriptions. Studies 5 and 9 did not cost any money, however, the cost saving came at the expense of the time taken to complete of the image description task. The descriptions produced by UK students (Study 6) and MTurk workers (Study 7) were produced in seven and five days respectively, whereas, the descriptions produced by Arab students and Crowdcrafting members were took approximately four weeks each to produce.

The overall quality of the descriptions was comparable between UK students who were paid, crowd members who were paid, and volunteer crowd members. Thus the Crowdsourcer should think of the cost in terms of money and time to collect the descriptions. Asking volunteer crowd members (such as Crowdcrafting) to undertake the task would save money, however, the time needed to complete the full task would probably be much longer than if the crowd members are paid. If time is a concern, then it is probably easier, faster and cheaper to crowdsource the image description task to MTurk workers, an approach which is successfully used by the VizWiz application (Bigham, et al. 2010). This would maintain a steady pool of crowd members whenever needed for image description.

Finally, it should be mentioned that none of these calculations have included the time and cost of setting up a crowdsourcing project, whether with volunteer students or paying crowd members. From the experience of this programme of research, this definitely takes time, but there is no particular difference between the two approaches in terms of set up and monitoring time and effort.

Table 10.3 Measures of costs in Studies 2 to 9

Measures	Study 2		Study 3		Study 4	Study 5		Study 6	Study 7	Study 9
	InMot	ExMot	InMot	ExMot	InMot	InMot	ExMot	ExMot	ExMot	ExMot
Payment per description (GBP)	None	None	None	None	None	None	None	0.20	0.08	None
Total cost (GBP)	0	0	0	0	0	0	0	80	82	0
Collection Period (days)	30	30	45	45	45	30	30	7	5	45

7. Participation in the TagIT Project - Making the Task Easier

Making the description task easier in Study 8 by asking students to tag objects in the images influenced students' participation rate, in that higher participation rate was obtained. The higher participation rate in the TagIT project in comparison to the DescribeIT project was found both with students and the crowd members of the CrowdCrafting.org platform (Study 9). In fact, the participation rate of only postgraduate students in UK in the TagIT project (34% participation rate for postgraduate students, 24% overall participation rate) was higher than the highest participation rate to the DescribeIT project, which was found with the Arab students (30% overall participation rate, this was a mixture of undergraduate and postgraduate students). Therefore, making the task less difficult would potentially motivate more people to participate in a socially responsible crowdsourcing projects.

The tags produced by both students and Crowdcrafting members were high in quality. The tags produced by students could potentially be easier to build on them to create image descriptions than those created by the Crowdcrafting members, as the students tended to create short sentences when tagging the images rather than single words, as was expected.

8. Quality Control and Training to Improve Quality and Engagement

The results of Study 10 showed that students in the quality control by peers condition (PeerQC) and the expert quality condition (ExpertQC) produced longer descriptions in comparison to the no quality control condition (NoQC) condition. However, unfortunately the Quality Control instructions did not affect the quality of the descriptions. Nonetheless, students in Study 10 were not provided with a feedback on their descriptions (from peers or experts) during the practical session, hence the effect of the Quality Control instructions was not examined on subsequent image descriptions.

On the other hand, a face-to-face training session with an expert did improve the image description quality in comparison providing only a set of guidelines and example of a typical image description. Although these points need further investigation, it seems that it would be better to put effort into training, perhaps by making a training video by an expert than putting effort into creating quality control mechanisms for the image descriptions.

9. Effect of Attitudes towards People with Disabilities and Altruism on Students' Motivation to Participate in the Image Description Task

The overall results suggest that there was no relation between the number of images produced per participant and their attitudes toward people with disability (as measured by the IDP sub-scales) and their general altruistic attitudes (SRA scale). However, comparing the scores of SRA and IDP scales obtained in this programme of research to previous research (See Table 10.4) yields interesting results.

Undergraduate students who participated in Phase 3 of Study 3 (the sample of non-participating students) and Study 10 were particularly low in altruism in comparison to the benchmark figures from Rushton et al. (1981). While this may have had an impact on students' lack of participation in Phase 2 of Study 3, the impact in Study 10 was not clear as students were supposed to be present for the practical session. However, interestingly Studies 5, 7, and 9 (TagIT Project) had the highest number of participants, and participants were particularly high in altruism in comparison to the figures in Rushton et al. (1981).

Surprisingly, the overall results of the studies conducted in this programme of research indicate that participants had a greater discomfort in social interactions with people with disabilities in comparison with those who participated in the study by Brown et al. (2009).

Table 10.4 Comparison of scores on SRA and IDP scales obtained in this programme of research to previous research studies

Study	IDP scores in comparison to Brown et al. (2009)	SRA scores in comparison to Rushton et al. (1981)
Study 3	significantly higher	significantly lower
Study 5	significantly higher	significantly higher
Study 6	n.s	n.s
Study 7	significantly higher	significantly higher
Study 8	significantly higher	n.s
Study 9 (TagIT Project)	significantly higher	significantly higher
Study 9 (DescribeIT Project)	significantly higher	n.s
Study 10	significantly higher	significantly lower

10.2 Original Contributions of the Research

Despite the increase in the number of initiatives to support disabled people by crowdsourcing projects, to the best of my knowledge, only one of the projects discussed in Chapter 2 (Section 2.3) investigated the motivations of the crowd, and reported only brief qualitative results. The series of empirical studies reported in this thesis contributes to our general understanding of what motivates people to participate in socially responsible crowdsourcing projects. To the best of my knowledge, this is the first series of empirical studies investigating which intrinsic and extrinsic motivational factors could motivate people to participate in socially responsible crowdsourcing projects to support disabled people. This is of importance because without understanding what drives people to participate in social responsible crowdsourcing projects, croudsourcers will be guessing what motivational factors could motivate their crowds or imitating motivational factors of other projects without understanding whether those motivational factors will work or not.

A second contribution that the series of studies reported in this thesis offers, is that contrary to previous research on the motivation of crowdsourcing projects which has relied on self-report of motivations, this research compared participants' self- reports of their motivations with their actual behaviour. As Antin and Shaw (2012) have noted, self-reports are vulnerable to social desirability bias, a fact borne out by the studies in this thesis. The studies in this thesis have shown that it is very important to study actual behaviour as well as self-reports, as these can be very different.

A third contribution of this thesis, is the detailed study of the quality of image descriptions produced for blind and partially sighted people, which compared a number of methods. In particular, to the best of my knowledge it is the first to consider a multi-dimensional method including Signal Detection Theory to measure the quality of image descriptions.

A fourth contribution is the investigation into cross-cultural differences on what would motivate people to participate in a socially responsible crowdsourcing project. The results suggested that there are differences between what motivated UK and Arab students to participate in a crowdsourcing project to support blind and partially sighted students. However, it should be noted that the UK students were a mixture of British and international students, and only one study was conducted with Arab students. Hence further research is needed to confirm whether students with different cultural backgrounds have different motivations to participate in socially responsible projects.

Finally, this thesis has implications for crowdsourcers to consider when creating projects for socially responsible tasks. Firstly, in order to motivate potential crowd members to participate in a socially responsible crowdsourcing project, crowdsourcers should consider making the task easy. This can be done by breaking difficult and off-putting tasks into smaller tasks, which make them easier to work on. This is likely to make people more motivated to participate. In the task proposed in this research asking participants to tag objects in an image in comparison to asking them to describe the whole image increased the participation rate. Secondly, for better quality work, crowdsourcer should consider training the crowd members to do a specific task rather than just providing guidelines or instructions on how to complete the task. As for the image description task proposed in this thesis training the students to do the task significantly improved the quality of the descriptions in comparison to only providing them with set of guidelines and an example. In the context of crowsourcing, training could be delivered via an engaging video rather than face-to-face, as was used in this research.

10.3 Validity of the Research Findings

Measuring motivation is always a challenging undertaking, because as Toure-Tillery and Fishbach (2014) noted it cannot be observed or directly recorded. Given the difficulty of measuring motivation, the studies conducted in this programme of research assessed the motivation of participants in the DescribeIT project by employing multiple measures including: The Situational Motivation Scale (SIMS), the number of participants who participated in particular versions of the project, the number of images described/tagged per participant, and the quality of the descriptions. In each condition of the studies a particular intrinsic or extrinsic motivational factor was investigated to understand its effects. However, human beings' motivations are complex, as we are often driven by a mix of motivations, but to understand how particular motivation affect participants in this programme of research only one intrinsic and extrinsic motivation was manipulated at a time.

Most of the participants recruited in this programme of research were students at the University of York in the UK. One exception was in Study 5 for which Arab students were recruited from Omar Al-Mokhtar University and Benghazi University in Libya and King Saud University in Saudi Arabia. The other exception was in Studies 7 and 9, in which participants were crowd members of established crowdsourcing platforms, Amazon Mechanical Turk and Crowdcrafting. The students recruited both from Arab counties and UK were not limited to a single discipline, instead they were deliberately recruited from diverse disciplines. In most of the studies students were asked to describe images from PowerPoint presentations which were related to modules they were studying, with some exceptions in Studies 5 and 6. However, in these studies and the one conducted with the committed crowd members (from Mechanical Turk and Crowdcrafting) the images used did not require a specific knowledge to understand.

The studies reported in Chapter 8 which investigated making the task simpler by using tagging rather than full description of the images were conducted with both students and committed crowd members to strengthen the validity and generalizability of the results obtained.

All studies except for Study 10 were run in ecologically valid, real-word settings. Study 10 was conducted as part of a practical session that students were undertaking as part of their module, which could limit the external validity of the results obtained in this study. However, this should not affect the internal validity of the results as students were randomly assigned to conditions.

10.4 Limitations of the Research

The research contributions discussed above need to be considered in light of the limitations of the studies conducted in this programme of research, which might have affected the validity of the findings from the studies.

The first limitation is based on the fact that Deci and Ryan's Self Determination Theory (SDT), which was the initial theoretical basis for the research, assumes three basic human needs that motivate people to initiate a behaviour, which are: the need for autonomy, competence and relatedness. Although support of these three needs was implemented in the DescribeIT project (See Chapter 3, section 3.2), they were not explicitly measured. However, SDT proposes that a higher level of intrinsic motivation and identified regulation are positively associated with perceived competence and autonomy (Deci and Ryan, 1985). Hence, the results obtained from Studies 5, 7, and 9 in regard to the level of intrinsic motivation and identified regulation as measured by the SIMS sub-scales Intrinsic Motivation and Identified Regulation (both above the midpoint of the scale), suggest that participants perceived competence and autonomy when participated in the DescribeIT project.

Another limitation is that the qualitative data analysis (content analysis) of what students said would motivate them and of other students may be subject to researcher bias and misunderstanding. However, this was mitigated by asking two individuals to work independently to code the statements from students and then work together till they reach an agreement on statements about each motivational factor. In addition, the data collected in the self-report phases is subject to social desirability bias. However, this was also mitigated by employing data triangulation, such as collecting both qualitative and quantitative data for analysis. But the most important strategy was the comparison of self-reports of motivation to actual behaviour.

The DescribeIT project used in Studies 2 to 5 was available for students to participate in for a period of four weeks, this may have been a short period of time to build up a crowd for a crowdsourcing project. However, due to the time constraints of conducting a programme of research studies to investigate the effectiveness of numerous intrinsic and extrinsic motivational factors, it was only possible to run each study for periods of four weeks.

In Study 2 the self-report phase showed that students were neutral in their thoughts about the extrinsic motivational factor improving own skills, however, this motivational factor was chosen for use in the actual behaviour phase for that study for several reasons: (1) Students were not significantly motivated by any of the extrinsic motivational factors presented in the self-report phase of the study; (2) Monetary reward (the highest rated and most mentioned) was not used, as that stage in the research I was interested to investigate non-financial motivational factors, and wanted to leave a study with monetary rewards until after most of the research with non-monetary rewards had been completed; (3) Getting academic credits was also not possible as a motivation, as it was not possible to reward students with academic credits. Thus, improving one's own skills was chosen as it was the highest rating among the extrinsic motivational factors presented.

There was a gender imbalance in the number of participants in the studies conducted. However, this programme of research did not investigate gender differences in what would motivate people to participate in a socially responsible crowdsourcing study, although this would be another interesting area of research. In addition, the University of York is a very multi-national university and the students who participated in Studies 2 to 4 were from a number of different countries, as has been mentioned in the relevant chapter there may have been some cross-cultural differences hidden in the data.

Another interesting limitation to consider in relation to Study 5 (with the Libyan and Saudi students), is the fact that the researcher was an Arab Libyan student herself, which might had introduced possibilities for inflation in the participation rate. Students, particularly those from Libya, might have wanted to participate in the DescribeIT project to support the researcher herself, or wanted to support research which promoted Libya.

Lastly, although the Situational Motivation Scale (SIMS) was proposed for use both in the field and laboratory settings, it was only tested by Guay, Vallerand, and Blanchard (2000) in three contexts, namely: education, interpersonal relationship, and leisure (sport). But in this programme of research it was used in a different context, that of a crowdsourcing task. In addition, for the purposes of this programme of research, some of the items of the SRA scale were thought to be inappropriate for the target populations of students in the UK, hence, four items from the scale were dropped, and when comparisons were made with previous research, scores were scaled appropriately. However, to make sure that the results obtained in this programme of research comparable to each other the same adapted SRA scale was used in all studies. The comparison of the SRA scores in the current set of studies with results of Rushton et al. (1981) is not ideal, as Rushton et al. (1981) results were in a different country from those participated in this research, and their data is quite old. In addition, the comparison of the IDP scores in the current set of studies with students in Brown et al. (2009) study is not ideal, as participants in the Brown et al study were occupational therapy students, who might be more comfortable in social interactions with people with disabilities than other students, considering they are expecting to work with disabled people during their careers.

10.5 Directions for future research

It is hoped that this programme of research has set the stage for further investigation to increase our understanding of what would motivate people to participate in socially responsible crowdsourcing projects. One of the promising research directions that future research can focus on is to examine the relation between giving participants feedback on their descriptions from lecturers or experts and the quality of the image descriptions produced. A loop of interaction between students and lecturers or experts by providing descriptions and getting feedback on them could potentially increase students' participation and engagement with the task.

In addition, future research should investigate whether there are further cross-cultural differences in what would motivate people to participate in crowdsourcing, and in particular to socially responsible crowdsourcing projects. There needs to be further exploration and validation of the results on cross-cultural differences obtained in Study 5.

It would be interesting for future research to investigate students' likelihood of participation in light of Theory of Planned Behaviour (TPB) (Armitage and Conner, 2001; George, 2004; Ajjan and Hartshorne, 2008; Pelling and Whitehas, 2009) been widely and successfully used to predict and explain behaviour in a wide range of domains. This could give a different perspective on students' motivations to participate in socially responsible projects.

It would also be interesting to see how self-reported and successful motivational factors can affect students' participation rate in the long run for sustained participation in socially responsible crowdsourcing projects.

Another interesting research direction is to try to produce initial image descriptions by using the tagging system with one set of students and then asking another set of students to improve these descriptions. It would be very interesting to see if students' participation rate would be higher if they build up their descriptions based on the tags produced by other students. Another possibility would be to use AI techniques to generate initial descriptions of the images, or at least the objects in the images and then ask students to correct and refine those initial descriptions.

A valuable extension to the methods used to measure the quality of the image descriptions in this programme of research is to consider the complexity or difficulty of the images in relation to the quality of the image descriptions.

Finally, it would be of great benefit to blind and partially sighted students to create a centralised database of image descriptions of common images used in education, such as standard diagrams, geographical maps and works of art. It could initially be based on images from frequently used textbooks in higher and further education. Such a database would then be accessible to all lecturers and blind and partially sighted students when needed.

10.6 Conclusions

Considering the overall results of this programme of research, it is clear that students in the UK were motivated by monetary rewards to participate in the DescribeIT project. They produced the highest number of image descriptions per student in comparison to other motivational factors. The quality of their descriptions was average overall and comparable to the quality of the descriptions of non-student participants (MTurkers and participants in the citizen science platform). The results also showed that postgraduate students were more motivated than undergraduate students to participate in the project to support blind and partially sighted students.

In addition, the overall results indicated that there could be a cultural factor, which was not controlled for in the studies conducted in the UK (which included students from many countries), that had an effect on students' participation rate. The participation rate of Arab students was much higher than of students in the UK, although there were a number of differences between the circumstances of the Arab students and those in the UK. Further investigation is needed to confirm these findings and the findings collected from the study with Arab students.

The face-to-face training with an expert improved the quality of the image descriptions in comparison to only providing a set of guidelines and example of a typical image description, although it may have also changed the motivational context. Furthermore, breaking down the image description task into a simpler tagging task also increased participation rate. The quality of the tags produced by students was high, and could potentially be used to build on to create image descriptions. Lastly, the results showed that there was no relation between the number of images produced per participant and their attitudes towards people with disability and their general altruistic attitudes.

At the end of this programme of research it was clear that investigating motivations is more complex than was anticipated at the beginning. Despite the large corpus of data collected in this programme of research, there are many other factors that were not systematically investigated that might have influenced students' motivation to participate in the DescribeIT project. For example, students' familiarity with the tasks they were invited to undertake. Students are probably more familiar with participating in an online questionnaire than describing images for a blind or partially sighted person. This fact could have influenced the large differences in students' participation rates in Phases 1 and 2 of the studies, with far higher participation rates in the online surveys in Phase 1 than in describing images in Phase 2.

In addition, the confidence of students to undertake the image description task might have influenced their participation. As some students noted (in the follow up interview of Study 3), they were not confident that they had the necessary skills to describe images for blind and partially sighted student (despite the guidelines provided). Accordingly, they decided not to submit their descriptions.

Other factors that could have influenced students' participation or lack of participation to the DesribeIT project might have been students' personality traits. For instant, students who are more open towards new situations and making new experiences would probably be more willing than others to participate in crowdsourcing projects even though they have not done so previously. Thus, if these traits were measured, we could probably have understood students' thoughts and behaviours somewhat better. Investigation of the influence of personality traits on students' participation to socially responsible crowdsourcing projects is a fruitful future topic of research.

Appendix A
Chapter 3

A.1. Guidelines for Describing Images for Blind and Partially Sighted Students:

1. Read the text provided with each image, the text can help you to understand the image, and you can build your description on that.

2. Think about the information that you can see in the image that is not available on the text on the slide. Think about what is important in the image and not discussed in the text.

3. In your description you should take the following aspects in your consideration:

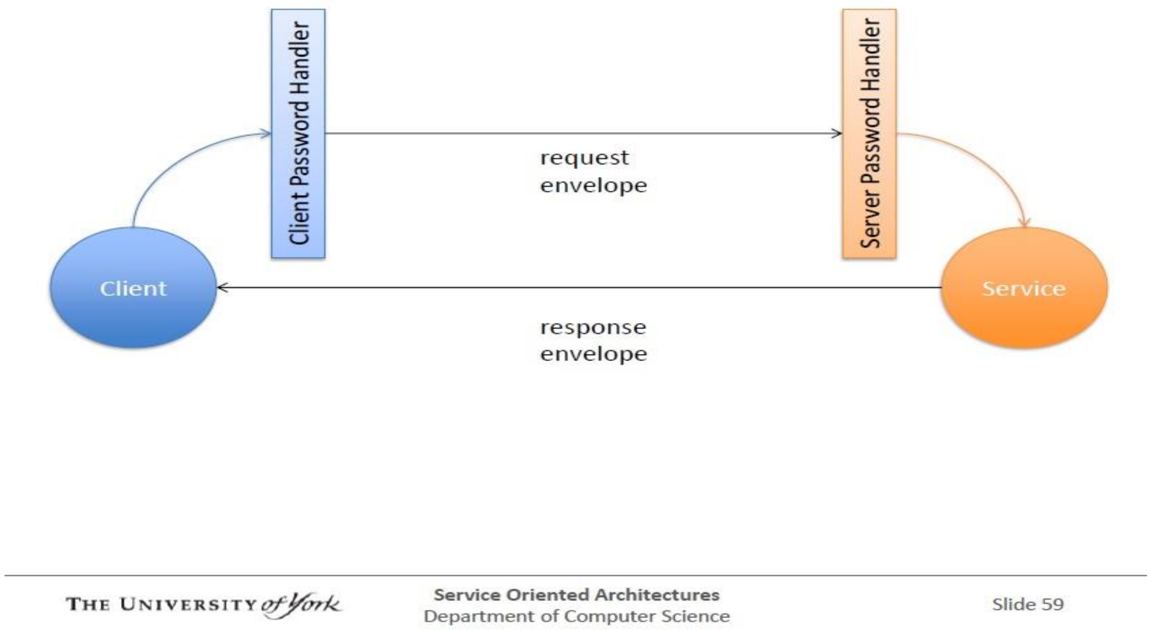
- What objects, people, locations are shown in the image?
- What are the interesting and important features of the contents of the image?
- What are the colours in the image – describing the colours in an image is very useful to blind students, even if they have never seen colour themselves, they usually know the colours of things, and describing colours can help them to learn about them.
- What are the sizes, orientations and relationships of objects – this information can be very helpful for blind students. So in your description think about the size and orientation of objects and describe objects in relation to one another.

4. Description length:

- Short description – the short description is often a one sentence description, which will provide blind and partially sighted students with an initial idea about what is in the image. In the short description you should describe the objects in the image, their basic features and shapes, any locations. Also include basic colours, particularly if they are important to the overall image.
- Long description – the long description provides more details and can be a number of sentences, depending on the complexity of the image. In the long description you can start with an overview (which might be the short description) and then go into much more detail. Try to take a logical approach: you might describe the image from top to bottom, going around in a clockwise

direction or start with the most important object in an image and then move on to less important objects. Provide as much detail as you think a student will need to understand the information contained in the image. That might be two sentences, it might be ten sentences, it's up to you.

Here is an example of an image description to help you (*this is not necessarily a perfect description, but will give you an idea of what might be useful*)



Short description:

A sequence of actions starting from a client, to client password handler, to server password handler, to service and back to client

Long description:

A coloured illustration which shows a sequence of actions. It consists of four main parts. On the left is a blue circle labelled "client", a curved arrow comes out of the top of the circle to reach a blue rectangle labelled client "password handler". A straight arrow labelled "request envelope" comes from the blue rectangle labelled "client password handler" to reach an orange rectangle labelled "server password handler". A curved arrow comes out of the left side of the orange rectangle to reach the top of an orange circle labelled "service". Lastly there is an arrow from the orange circle labelled "service" to the blue circle labelled "client", the arrow is labelled "response envelope"

Appendix B
Chapter 4

B.1. Study 1

B.1.1 Information Sheet and Informed Consent

THE UNIVERSITY of York

Information Sheet and Informed Consent

1. Who is running the study?

The study is being run by Fatma Layas, a research student in the Department of Computer Science at the University of York. Fatma is supervised by Prof Helen Petrie who is a Professor of Human Computer Interaction and Dr Christopher Power, who is a Lecturer at the Department of Computer Science, the University of York.

2. What is the purpose of the study?

The study aims to explore students' motivations in using an application to provide descriptions of PowerPoint images for their blind and partially sighted peers.

3. What will I have to do?

You will be asked simple set of questions about your motivations to use an application to make image descriptions for blind and partially sighted students. You will also be asked basic personal information. The study ends after you have submitted your answers, and you should receive your Amazon voucher gift code within 48 hours of your submission.

4. Who will have access to the study data?

Fatma Layas, and her supervisors' Helen Petrie and Christopher Power will have access to the study data.

5. What will happen to the information which I give?

The data will be kept confidential for the duration of the researcher study. On completion of the researcher study, they will be retained for a further seven years and then destroyed.

6. Will my participation in the study be kept confidential?

Your anonymity will be maintained at all times and no comments will be ascribed to you by name in any written document. Nor will any data be used from the study that might identify you individually. Any extracts from what you will write that are quoted in the study's report will be entirely anonymous.

7. Do I have to take part of the study?

Your participation in the study is completely voluntary. You will be free to withdraw from the study at any time and/or request that your transcript not be used. You will not be penalised for withdrawing nor will you be questioned on why you have withdrawn.

8. Still have concerns?

If you have any queries concerning the nature of the research or are unclear about the extent of your involvement in it, please contact me, Fatma Layas, at fal503@york.ac.uk.

If you are agreeing to take part in the study under the conditions mentioned above, please tick the appropriate boxes.

I, the undersigned, confirm that (please tick box as appropriate):

1.	I have read and understood the information about the project, as provided in the Information Sheet.	☐
2.	I have been given the opportunity to ask questions about the project and my participation.	☐
3.	I voluntarily agree to participate in the project.	☐
4.	I understand I can withdraw at any time without giving reasons and that I will not be penalised for withdrawing nor will I be questioned on why I have withdrawn.	☐
5.	The procedures regarding confidentiality have been clearly explained to me.	☐
6.	The use of the data in research and sharing has been explained to me.	☐
7.	I, agree to sign and date this informed consent form.	☐

Participants' signature:

Researcher: Fatma Layas (fal503@york.ac.uk)
Supervisors: Helen Petrie (helen.petrie@york.ac.uk)
Christopher Power (christopher.power@york.ac.uk)

B.1.2. Study 1 Recruitment Email

Dear student,

My name is Fatma Layas, I am a PhD student in the Department of Computer Science supervised by Professor Helen Petrie. My PhD research is about using crowdsourcing to help blind and partially sighted students.

As students, we use many electronic teaching materials, such as Powerpoint presentations, on a daily basis. Blind and partially sighted students can access these materials using software called screenreaders. Screenreaders convert the text in electronic teaching materials into speech so that blind and partially sighted students can listen to the material. However, screenreaders cannot provide any information about images, and electronic teaching materials contain many images. The only way that a screenreader can convey the information in an image is having someone create a description of the image which can then be read out.

My research is about creating an online crowdsourcing application that will allow sighted students to very easily add descriptions to images on the electronic teaching materials for their courses, such as the PowerPoint slides that the lecturers use. Students would be able to describe as many or as few images as they wanted to. The system will include instruction on how to create good image descriptions.

At the moment I am interested to find out what would motivate sighted students to participate in such a project. The questionnaire asks a simple set of questions about your thoughts about the project, you should find it interesting and it should take no longer than 15 minutes to complete.

Your responses will be completely confidential and anonymous. Any information about the results from the questionnaire will only be reported in ways in individuals cannot be identified.

Everyone who completes the questionnaire will be entered into a prize draw for one of 10 Amazon gift vouchers, each valued at £10.

If you are interested in helping, the questionnaire can be accessed here:

[***A link to the questionnaire on QuestionPro.com**]

I hope that you will be able to help me with my research. If you have any questions, please contact me at fal503@york.ac.uk.

Thank you for taking the time to consider my request.

Fatma Layas, Phd student

and

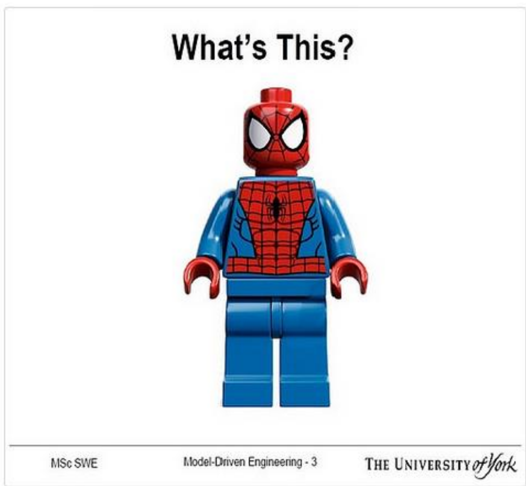
Helen Petrie Phd AFBPsS CPsychol FRSA

B.1.3. Online Questionnaire

Thank you for offering to take part in my research, which is about students' attitudes to participating in a crowdsourcing project. The project is to create descriptions of the images in electronic teaching materials such as PowerPoint slides. This will allow blind and partially sighted students to understand the images in the materials, which they would not be able to do otherwise. The questionnaire should take about 15 minutes to complete. Hopefully you will find it interesting and it will greatly help our research. Your responses will be completely confidential and anonymous. Any information about the results will only be reported in ways in which individuals cannot be identified. Everyone who completes the questionnaire will be entered into a prize draw for one of 10 Amazon gift vouchers, each valued at £10. If you have any questions, please contact email me, Fatma Layas, at fal503@york.ac.uk.

Introduction

Electronic teaching materials, such as PowerPoint presentations, contain many images. Many of these images are vital to understanding the materials and being able to learn from them. Blind and partially sighted students may not be able to learn effectively if they do not understand what is in the images, but unfortunately lecturers do not provide explicit descriptions of every image they use in their presentations for a variety of reasons. I am developing a crowdsourcing project to allow sighted students to provide descriptions of the images in electronic teaching materials on their courses for blind and partially sighted students. The project will allow students to describe as many images or as few images as they wish. Students will be provided with guidelines of how to describe images most effectively for blind and partially sighted students. The following image shows the possible layout of the application for the project. When you have finished reviewing the layout of the application, please select the Continue Button



Your views on the image description project

How likely is it that you would participate in this project?

Not at all likely to participate ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very likely to participate

What would motivate you to participate in this project?

How likely is it that other students on your course would participate in this project?

Not at all likely to participate ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very likely to participate

What do you think would motivate other students on your course to participate in this project?

Please answer the questions below about whether these factors would motivate you to participate in the image description project

Your sense of altruism, wanting to help other students

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Improving your academic skills

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Being paid for your efforts

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Enhancing your job opportunities in the future

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Being connected with other students on your course

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

The fun and entertainment of the activity

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

To pass the time

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Knowing that you are contributing to a large project

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

The social recognition you would receive

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Drawing attention to your skills

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Being in a competition with other students

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Getting academic credits

Not at all ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very much

Why did you give that particular rating?

Do you think there are any other factors which would motivate you to participate in the image description project? If so, please describe briefly.

Finally, for statistical purposes, a few questions about yourself

Are you:
1. Male
2. Female

What is your age?

What country are you from?

In which department are you studying?
1. Archaeology
2. Biology
3. Computer Science
4. English
5. Health sciences
6. History
7. Psychology
8. Theatre, Film and Television
9. Other

What degree are you studying for?
1. Diploma/Certificate
2. Undergraduate (BA, BSc, BEng etc)
3. Taught postgraduate (MA, MSc etc)
4. Research degree (research MA/MSc, PhD)
5. Other

Do you use social media sites?

Never ❑ ❑ ❑ ❑ ❑ ❑ ❑ Very regularly

Which social media do you use (Select all that apply)?
1. Facebook
2. Twitter
3. LinkedIn
4. Pinterest
5. Google Plus+
6. Instagram
7. Other

Have you participated in any crowdsourcing projects?
Crowdsourcing is using the power of the Web to recruit a large number of people to undertake activities, often dividing the activity into small, easy to complete parts. Wikipedia is a famous example of crowdsourcing project.

1. Yes
2. No

In which crowdsourcing project/s have you participated?

Why did you participate in these crowdsourcing projects?

If you would like to be entered into the prize draw for one of 10 Amazon gift vouchers, each valued at 10£, please provide your University email address. This will not be used for any other purpose, and will be destroyed as soon as we make the prize draw.

Appendix C
Chapter 5

C.1. Study 2

C.1.1. Phase 1: Recruitment email

The same email used in Phase 1 Study 1, See B.1.2

C.1.2. Phase 2: Recruitment emails

C.1.2.1. ExMot condition

Title: Help with piloting DescribeIT and improve your skills!

Dear Students,

We are piloting the "DescribeIT" project; an online system to support blind and partially sighted students taking the department's numerous programmes. We are asking you to help us in a small way with this pilot study.

The project is about improving access to teaching materials for blind and partially sighted students. In particular, to improve the access to images that are part of the PowerPoint slides that lecturers use in teaching. The study involves evaluating the system we are designing to help students describe these images.

PowerPoint presentations used in teaching contain many images. Some of thes e images are only for decoration. However, most images are important to understanding the material and being able to learn it. Blind and partially sighted students are not able to learn effectively if they do not know what is in the images, but lecturers u nfortunately do not have time to provide explicit explanations of every image they use in their presentations that will make the material accessible to these students.

The DescribeIT project will address this problem by allowing sighted students to provide descriptions of images in presentations for their blind and partially sighted fellow students. . The project will allow students to describe as many or as few images as they wish at any time.

Participating in this project will improve your skills to make websites, apps and software accessible to people in the future. Following the image description guidelines provided with this system will help you learn how to provide good image descriptions. In addition, this can help you improve your skills in how to co mmunicate clearly in computer science using images.

I would be very grateful if you would be willing to take part in our study. If you are interested, please use the link below to access the project DescribeIT:

[*****The project link******]

Many thanks

C.1.2.2. InMot condition

Title: Help pilot DescribeIT and help blind and partially sighted students

Dear Students,

We are piloting an online project to support blind and partially sighted students taking the department's numerous programmes, it is called the DescribeIT. We are asking you to help us in a small way with this pilot study.

The project is about improving access to teaching materials for blind and partially sighted students. In particular, to improve access to the many images in PowerPoint slides that lecturers use for teaching. For our pilot study involves asking students like yourself to try out the system and describe some images using it.

PowerPoint presentations used in teaching contain many images. Some of these images are only for decoration. However, many images are important to understanding the material and being able to learn it. Blind and partially sighted students are not able to learn effectively if they do not know what is in the images, but unfortunately lecturers do not always have time to provide explicit explanations of every image that will make the material accessible to these students.

The DescribeIT project will address this problem by allowing sighted students to provide descriptions of images in presentations for their blind and partially sighted fellow students. The project will allow students to describe as many or as few images as they wish at any time.

The care you provide to others can only give you good things in return. Please help us improve the accessibility of images presented in PowerPoint slides.

I would be very grateful if you would be willing to take part in our study. If you are interested, please use the link below to access the project DescribeIT:

[******The project link********]

Kind regards

Helen

C.1.3. Phase 2: Online Questionnaire

Thank you for agreeing to take part in this survey about describing images for blind and partially sighted students. This survey should only take 5 - 7 minutes to complete. Everyone who completes the questionnaire will be entered into a prize draw for one of 10 Amazon gift vouchers, each valued at £10. If you have any questions, please contact email me, Fatma Layas, at fal503@york.ac.uk.

Please read each statement carefully. Using the scale below, please select the ch oice that best describes the reason why were you engaged in the image description activity?

	Not all	A very little	A little	Moderately	Enough	A lot	Exactly
Because I think that this activity is interesting	❑	❑	❑	❑	❑	❑	❑
Because I am doing it for my own good	❑	❑	❑	❑	❑	❑	❑
Because I am supposed to do it	❑	❑	❑	❑	❑	❑	❑
There may be good reasons to do this activity, but personally I don't see any	❑	❑	❑	❑	❑	❑	❑
Because I think that this activity is pleasant	❑	❑	❑	❑	❑	❑	❑
Because I think that this activity is good for me	❑	❑	❑	❑	❑	❑	❑
Because it is something that I have to do	❑	❑	❑	❑	❑	❑	❑
I do this activity but I am not sure if it is worth it	❑	❑	❑	❑	❑	❑	❑
Because this activity is fun	❑	❑	❑	❑	❑	❑	❑
By personal decision	❑	❑	❑	❑	❑	❑	❑
Because I don't have any choice	❑	❑	❑	❑	❑	❑	❑
I don't know; I don't see what this activity brings me	❑	❑	❑	❑	❑	❑	❑
Because I feel good when doing this activity	❑	❑	❑	❑	❑	❑	❑
Because I believe that this activity is important for me	❑	❑	❑	❑	❑	❑	❑
Because I feel that I have to do it	❑	❑	❑	❑	❑	❑	❑
I do this activity, but I am not sure it is a good thing to pursue it	❑	❑	❑	❑	❑	❑	❑

Select the category on the right that conforms to the frequency with which you have carried out the following acts.

	Never	Once	More than once	Often	Very often
I have given directions to a stranger.	❑	❑	❑	❑	❑
I have given money to a charity.	❑	❑	❑	❑	❑
I have given money to a stranger who needed it (or asked me for it).	❑	❑	❑	❑	❑
I have done volunteer work for a charity.	❑	❑	❑	❑	❑
I have donated blood.	❑	❑	❑	❑	❑
I have helped carry a stranger's belongings (books, parcels, etc.).	❑	❑	❑	❑	❑
I have delayed an elevator and held the door open for a stranger.	❑	❑	❑	❑	❑
I have allowed someone to go ahead of me in a lineup (at photocopy machine, in the supermarket).	❑	❑	❑	❑	❑
I have given a stranger a lift in my car.	❑	❑	❑	❑	❑
I have pointed out a clerk's error (in a bank, at the supermarket) in undercharging me for an item.	❑	❑	❑	❑	❑
I have let a neighbour whom I did not know too well borrow an item of some value to me (e.g., a dish, tools, etc.)	❑	❑	❑	❑	❑
I have bought 'charity" Christmas cards deliberately because I knew it was a good cause.	❑	❑	❑	❑	❑
I have helped a classmate who I did not know that well with a homework assignment when my knowledge was greater than his or hers.	❑	❑	❑	❑	❑
I have offered to help a handicapped or elderly stranger across a street.	❑	❑	❑	❑	❑
I have offered my seat on a bus or train to a stranger who was standing.	❑	❑	❑	❑	❑
I have helped an acquaintance to move households.	❑	❑	❑	❑	❑

We acknowledge that the way we respond to people depends on how well we know them as individuals, however, for the purpose of this study we would like to know how you feel in general when you meet a person with a disability. Please read each statement carefully and decide how much it describes how you feel.

	I disagree very much	I disagree somewhat	I disagree a little	I agree a little	I agree somewhat	I agree very much
It is rewarding when I am able to help	❑	❑	❑	❑	❑	❑
It hurts me when they want to do something and can"t	❑	❑	❑	❑	❑	❑
I feel frustrated because I don't know how to help	❑	❑	❑	❑	❑	❑
Contact with a person with a disability reminds me of my own vulnerability	❑	❑	❑	❑	❑	❑
I wonder how I would feel if I had this disability	❑	❑	❑	❑	❑	❑
I feel ignorant about people with disabilities	❑	❑	❑	❑	❑	❑
I am grateful that I do not have such a burden	❑	❑	❑	❑	❑	❑
I try to act normally and ignore the disability	❑	❑	❑	❑	❑	❑
I feel uncomfortable and find it hard to relax	❑	❑	❑	❑	❑	❑
I am aware of the problems that people with disabilities face	❑	❑	❑	❑	❑	❑
I can't help staring at them	❑	❑	❑	❑	❑	❑
I feel unsure because I don't know how to behave	❑	❑	❑	❑	❑	❑
I admire their ability to cope	❑	❑	❑	❑	❑	❑
I don't pity them	❑	❑	❑	❑	❑	❑
After frequent contact, I find I just notice the person not the disability	❑	❑	❑	❑	❑	❑
I feel overwhelmed with discomfort about my lack of disability	❑	❑	❑	❑	❑	❑
I am afraid to look at the person straight in the face	❑	❑	❑	❑	❑	❑
I tend to make contacts only brief and finish them as quickly as possible	❑	❑	❑	❑	❑	❑
I feel better with people with disabilities after I have discussed their disability with them	❑	❑	❑	❑	❑	❑
I dread the thought that I could eventually end up like them	❑	❑	❑	❑	❑	❑

Finally, for statistical purposes, a few questions about yourself [Same as in B.1.3.]

C.1.4. Examples of Students' Comments

Table C.1 Comments from students about sense of altruism and wanting to help others

Likert Ratings	Students' comments
1	I volunteer with disability sport, there's only so much I can do. (P16)
4	Because of the immense work, I have to do to help myself in order to complete my studies. (P19)
5	I am willing to help others, but I may not spare time to do this when I have many work to do. (P20)
6	I would happily contribute to making some other student's life easier. Especially if the contribution is as small as describing an image. (P12)
7	help others who need my help always makes me happy (P13)

Table C.2 Comments from students about being paid for their efforts

Likert Ratings	Students' comments
1	I'm not the kind of person that does things for money. And I wouldn't do this as a full-time job - too boring. (P21)
4	As I said before, because of my tight schedule reward is important. If I had more free time I would consider participating to this activity only for the cause. (P5)
5	It is good to have some extra (P8)
6	Adding a financial award would provide a further incentive that could attract participants that wouldn't otherwise take part. (P17)
7	I may be altruistic, but I would like at least some compensation, even in the form of credits or score not necessarily money. Mainly so I feel like I am not being employed to do it. Also, compensation makes me want to do it better than for free. (P6)

Table C.3 Comments from students about knowing that they contribute to a large project

Likert Ratings	Students’ comments
2	I'd be doing it for more intrinsic reasons. (P16)
3	Not a main reason (P13)
4	I don’t mind (P14)
5	If something is going to be helpful for a bigger crowd; as in my contribution helping more people out, sounds like a good incentive too. (P7)
6	I would agree with this if the project is of benefit to other people (P4)
7	I feel it is a honor [honour] and I will feel proud of my contribution. (P20)

Table C.4 Comments from students about getting academic credits

Likert Ratings	Students’ comments
1	I think academic credits should be obtained as a result of your academic performance in a particular module or in a research project. I do not consider this activity a way to demonstrate my academic performance. (P3)
3	If you get credits is a plus but for me is not the mainly purpose (P22)
4	Not a big deal. Grades depend on how much you study (P11)
6	Anything that excels career issues are always welcoming (P10)
7	Extra credits are always welcome. The same motivation would exist if I knew that I would lose credits for a course if I didn't complete the task. (P12)

Table C.5 Comments from students about improving their academic skills

Likert Ratings	Students' comments
1	I don't see how this would improve my skills. (P12)
2	I don't think I would see a significant improvement on my academic skills by doing this activity. (P19)
4	I don't think this particular activity would improve my skills. However describing images may help improve my knowledge by focusing my attention on how to better explain the presentation content to others. (P3)
5	Building yourself is always good (P8)
6	Because I want to have an expertise on my academic skills and if also this help in somewat [somewhat] to the society it could be exceptional (P22)
7	I love learning about things, especially free. (P6)

Table C.6 Comments from students about the fun and entertainment of the activity

Likert Ratings	Students' comments
1	I wouldn't enjoy it. (P16)
2	Maybe it's not so fun especially on large and complex images. (P19)
3	I want to learn more things than enjoy the fun of the activity. (P20)
4	Not a big deal. (P11)
5	I think that if you do what you like even if it's hard or tired it could be fun (P22)
6	Fun part is of paramount of importance. Because if things are boring, I might drift away to something else and never come back. (P7)
7	Seems like fun. The idea of wracking my brains to describe something is intriguing (P6)

Table C.7 Comments from students about being connected with other students

Likert Ratings	Students' comments
1	I prefer other offline ways to socialize than through apps. (P3)
2	This wouldn't be an major incentive for me but it could be positive. (P17)
3	connecting with others have many ways. I do not feel it is very essential. (P20)
4	It would probably bring me closer to my classmates especially if there were any visual impaired. (P19)
5	It would positively enhance my relationship with the person having eyesight problems. (P12)
6	It is important to me to make connections and this could be another way to do so (P4)
7	Brilliant idea (P14)

Table C.8 Comments from students about enhancing job opportunities

Likert Ratings	Students' comments
1	I prefer to enhance my opportunities building networks, improving my knowledge and skills. This will not be a motivator for me. (P3)
2	This activity is not relevant with my field. (P5)
4	If job opportunity on the line, definitely it is a good incentive. (P7)
5	Good way to get / increase portfolio (P10)
6	job opportunities is very important for students. It may be a potential chance for me to find a job. (P20)

Table C.9 Comments from students about being in a competition with other students

Likert Ratings	Students' comments
1	I think it's a silly thing to compete over. (P21)
2	Is not a competion [competition], is trying to help (P22)
3	I do not think it is essential to compete with others in this activity. (P20)
4	I like to be in a team, not winning is the motto. However, making team better is - to me. (P7)
5	While this can increase productivity i believe it can compromise quality. (P4)
6	Standing out of the crowd (P8)
7	I would join the competition for sure. (P3)

Table C.10 Comments from students about drawing attention to their skills

Likert Ratings	Students' comments
1	Other factors are more important to me. (P3)
4	Not really fussed about it. If I am skilled, I know it. (P7)
5	Good to know which skills are getting renowned (P8)
7	I think it is good to learn something and enhance my skills. (P20)

Table C.11 Comments from students about social recognition they would receive

Likert Ratings	Students' comments
1	I don't care about social recognition. I believe people should do things for self-satisfaction and to help - not for other people to think highly of them. (P21)
2	This is not important to me as I find you should do something good for that reason, not because of rewards. (P4)
3	It would reflect good on me if others knew I cared about this. Or reversely, it would reflect bad on me if others knew I didn't care. (P12)
4	not bothered (P9)
5	It would be nice to have recognition for it. Maybe a weekly leaderboard would make me want to do more. (P6)
6	I am interested in understanding the society. (P20)

Table C.12 Comments from students about passing the time

Likert Ratings	Students' comments
1	Far better ways to pass time than giving descriptions of pictures :) (P21)
2	Maybe not. Time is valuable, I have better things to do. (P7)
3	I have more fun things if I want to just pass the time. (P6)
4	Not a big deal. You can always YouTube and learn things or Code~ (P11)
5	If it's fun, passing time sounds good (P8)
7	Definetely [Definitely]. I only mostly use apps when I have free time. Even if a I am too motivated or altruistic I have many things to do. If I have free time I will be using it. (P3)

C.2. Study 3

C.2.1. Phase 2: Recruitment email for ExMot condition

Email title: Help with piloting DescribeIT and improve your skills!

Email content:

Dear Students,

We are piloting the “DescribeIT” project; an online system to support blind and partially sighted students, by providing access to images that are part of the PowerPoint slides that lecturers use in teaching. *More information can be found in the project Info page*

The DescribeIT project allows sighted students to provide descriptions of images in presentations for their blind and partially sighted peers. Describing the images in the Programming for Digital Media course will increase your understanding of the course materials. In addition, it will help you improve your skills in how to communicate clearly in Programming for Digital Media using images.

I would be very grateful if you would be willing to take part in our study. If you are interested, please use the link below to access the project DescribeIT, you will need this password to access the project: *******

[****The Project Link*****]

Many thanks

Helen

Helen Petrie Phd AFBPsS CPsychol FRSA

C.2.2. Phase 2: Recruitment email for InMot condition

Email title: Help pilot DescribeIT and help blind and partially sighted students.

Email content:

Dear Students,

We are now piloting the “DescribeIT” project, an online system to support blind and partially sighted students by providing descriptions of the images in the PowerPoint slides that lecturers use in teaching.

As part of the pilot, we are asking you to describe some of the images from your HACS class. This can potentially help blind and partially sighted students who take the module in the future. To have an idea about what type of difficulty a blind and partially sighted students face when the teaching materials are not accessible and how this project will help them, please watch this short interview with Ian, a blind student from the University of Hull. [***link to the video***]

I would be very grateful if you would be willing to take part in this pilot study. If you are interested, please use the link below to access the project DescribeIT, you will need this password to access the project: *******

[****The Project Link*****]

Kind regards

Helen

Helen Petrie Phd AFBPsS CPsychol FRSA

C.2.3. Phase 3: Recruitment email

Dear students,

We are conducting interviews as part of a research study to increase our understanding of what would motivate students to help their blind and partially sighted peers. As a student you are in an ideal position to give us valuable first-hand information about this issue from your own perspective, so I writing to ask if you could do a short interview with me.

The interview should take no longer than 15 minutes and your responses to the questions will be kept confidential. As a thank you, you will be offered a £15 Amazon voucher for your time and effort. The interview will take a place in the computer science department, Heslington East, the specific room number will be sent to you once you choose your slot. I will be conducting interviews between February 1 and February 12 and I am very flexible in times I’m available.

If you are able to participate please choose your preferred date and time from the doodle below and I'll contact you to confirm your slot.

[****doodle link*****]

If you have any questions please do not hesitate to contact me at Fal503@york.ac.uk.

Thank you in advance!

C.2.4. Examples of Students' Comments:

Table C.13 Comments from students about altruism and wanting to help others

Likert Ratings	Students' comments
4	I like to help others when I have the time to help. (P2)
5	I try to help others but usually only when the opportunity lands at my feet, I am not proactive about it (P12)
6	People with sight issues should be helped in any way possible. This doesn't require a lot of effort form sighted students, so everybody wins! (P14)
7	Because helping people is important. If you can help someone, you are morally obligated to do so. (P21)

Table C.14 Comments from students about getting academic credits

Likert Ratings	Students' comments
1	I do not see this as a way to gain credit on my course. (P25)
3	Although it's nice, it's not hugely important to me. (P4)
5	Every little helps. (P15)
6	If it was possible, that would be great (P6)
7	Anything to improve my overall degree. (P3)

Table C.15 Comments from students about enhancing job opportunities

Likert Ratings	Students' comments
1	I see no relation between participating in crowd sourcing that one time and future job opportunities. (P5)
2	I think that the skill of more in depth explanations would beca[become] good skill to add to my CV. however I don't think volunteering for research would enhance my job opportunities unless I wanted to go into this area of research. (P2)
3	If it did enhance my job opportunities that would be great, but I don't see how it could (P6)
4	It's good to have on a CV to a certain extent, but I feel a larger role is important for it to be useful. (P18)
5	I am always looking for ways to enhance my CV. (P10)
6	I want to show that I've been productive and helpful while at university, and that I have attempted to better the environment for not only myself, but others too. (P3)
7	I want to be in the best possible career for the future and working with visually impaired people will always help when employers are looking at my skills (P15)

Table C.16 Comments from students about being paid for their efforts

Likert Ratings	Students' comments
1	Money should have no bearing on education. (P21)
2	Although a nice gesture, it is not that important to be paid. (P4)
3	At the end of the day, there are more important things than being paid (P11)
4	Depend on the pay. (P5)
5	If the money at least minimum wage I could treat it like a part time job (P6)
6	Being paid is naturally motivating but would also take away from the altruistic nature of the project. (P12)
7	Students live on a small budget, any income is more than welcome (P9)

Table C.17 Comments from students about the fun and entertainment of the activity

Likert Ratings	Students' comments
2	Its more about helping and being involved than it being fun and entertaining (P9)
3	may be interesting but not much of a motivation (P25)
4	Depends very heavily on the implementation, it could be fun to participate in, but it could also be very tedious and mechanical (P12)
5	I'm quite happy to sit and perform routine tasks as long as they are benefiting other people (P15)
6	The enjoyment is important for keeping people engaged. (P21)

Table C.18 Comments from students about being connected with other students

Likert Ratings	Students' comments
2	I feel I already have enough opportunities to connect with others on my course. (P10)
3	I do not feel too strongly about others on my course, I tend to socialise with people outside of it. (P12)
4	Cooperating on a abstract level is fun. (P5)
5	Making lots of friends on the course is always good as you can get help from many different people and expand your knowledge on subjects. (P15)
6	Being connected with other like-minded students allows people to develop ideas even further (P8)
7	I feel it is important to communicate and be close to peers as it means you can discuss problems. (P2)

Table C.19 Comments from students about improving academic skills

Likert Ratings	Students' comments
1	I don't think will help much in my personal academic skills (P25)
2	I do not feel that I would benefit much academically from this, if at all (P12)
3	While it does make you consider what is happening in an image in greater detail, there are better ways to improve my academic skills (P6)
4	I think it would help me think more in depth about the detail I give when describing objects and this could help me if I needed to explain things. (P2)
5	I am fairly motivated to gain as much experience in any way I can from my university experience. (P10)
6	Academic skills are very important and they will help you get a job. (P21)
7	I think participating would also benefit me since it requires me to describe images that are relevant to my course and would require me to develop my technical writing skill (P26)

Table C.20 Comments from students about knowing that you contributing to a large project

Likert Ratings	Students' comments
2	That just means I am knowing that my contribution will be small and count for very little (P20)
3	It helps to know that people will actually read what I write, but the scale of the project does not motivate me (P6)
4	There's a sense of involvement (P18)
5	Could be motivational if the project is large enough to make progress quickly, and if the real-world benefit could be felt (P12)
6	It makes a person feel important to feel like that are part of something and doing well for a lot of people, not just for themselves. I would include myself in this. (P4)
7	part of the main reason why I am taking part (P25)

Table C.21 Comments from students about drawing attention to skills

Likert Ratings	Students' comments
2	It does not seem a difficult task, the skill requirement is very low (P12)
4	I'd have to use it before knowing whether it'd help me discover any unknown skills (P14)
5	I really want to be able to improve my skills at every opportunity, and showcasing these holds no disadvantages as far as I can see. (P3)

Table C.22 Comments from students about being in competition with other students

Likert Ratings	Students' comments
1	Being in competition over helping someone who needs it, is not appealing at all to me. I feel if students are going to do it based on competition, then it is for the wrong reason completely. (P4)
2	I am adverse to being in competition with other students. (P10)
4	Competition with other students isn't really what inspires or motivates me (P8)
5	Always like a bit of competition (P20)
6	I'm competitive, and it would be a good way to get to know people (P6)
7	Competition is fun and usually leads to good results (P14)

Table C.23 Comments from students about passing the time

Likert Ratings	Students' comments
1	I barely have enough time on my hands as it is! (P3)
2	I already have many things that fill in my time (P4)
3	I have other means of filling my time if needed. (P10)
4	Everything exist to pass time at least a little. (P5)
5	It would be good for passing the time if I could do it on my mobile, but at home I would be more likely to play a game (P6)
6	I'm always looking for pastimes that don't require too much time (unless I want them to) (P14)

Table C.24 Comments from students about the social recognition they would receive

Likert Ratings	Students' comments
1	I don't feel that social recognition would be needed unless to [it] was towards encouraging other to help. (P2)
2	Although pleasant, not very important and not the true reason for why I would participate in the project. (P4)
3	I don't feel it's a project that would cause you to receive much social recognition (P18)
4	I am unsure how much recognition would be received, but this doesn't interest me as much anyway. (P3)
5	Although I'm not sure how it'd happen, it'd be nice to be recognised for something so positive (P14)

C.2.5 Information Sheet

THE UNIVERSITY of York

Information Sheet

1. Who is running the study?

The study is being run by Fatma Layas, a research student in the Department of Computer Science at the University of York. Fatma is supervised by Prof Helen Petrie who is a Professor of Human Computer Interaction and Dr Christopher Power, who is a Lecturer at the Department of Computer Science, the University of York.

2. What is the purpose of the study?

The study aims to explore students' motivations in using a project to provide descriptions of PowerPoint images for their blind and partially sighted peers.

3. What will I have to do?

You will be asked simple set of questions about your motivations to use a project to make image descriptions for blind and partially sighted students. You will also be asked basic personal information. The study ends after you have submitted your answers, and you should receive your Amazon voucher gift code within 48 hours of your submission.

4. Who will have access to the study data?

Fatma Layas, and her supervisors' Helen Petrie and Christopher Power will have access to the study data.

5. What will happen to the information which I give?

The data will be kept confidential for the duration of the researcher study. On completion of the researcher study, they will be retained for a further ten years and then destroyed.

6. Will my participation in the study be kept confidential?

Your anonymity will be maintained at all times and no comments will be ascribed to you by name in any written document. Nor will any data be used from the study that might identify you individually. Any extracts from what you will write that are quoted in the study's report will be entirely anonymous.

7. Do I have to take part of the study?

Your participation in the study is completely voluntary. You will be free to withdraw from the study at any time and/or request that your transcript not be used. You will not be penalised for withdrawing nor will you be questioned on why you have withdrawn.

8. Still have concerns?

If you have any queries concerning the nature of the research or are unclear about the extent of your involvement in it, please feel free to ask me or contact me, Fatma Layas, at fal503@york.ac.uk.

C.2.6 The DescribeIT Project

C.3. Study 4

C.3.1. Phase 1: Recruitment email

Hi all,

This is Helen Petrie, who spoke in the Using Primary Materials lecture this morning about our project to create a crowdsourcing system to describe images in teaching materials to help blind and partially sighted students.

If you have 10 minutes or so and can help our project by completing the questionnaire to gauge students' opinions about this project, here's the link to the questionnaire:

[**** Link to the online questionnaire****]

I should have mentioned this morning that all information in the questionnaire will be completely confidential and anonymous and we will only use email addresses to randomly choose the 10 people who will win the £10 Amazon gift vouchers.

If you weren't in the lecture this morning and you are wondering what all this is about, I've included a summary of what this research project is about below.

If you have any questions, feel free to email me.

Many thanks for your help.

Cheers

Helen

The project:

We now use many electronic materials in learning, such as PowerPoint presentations. Blind and partially sighted students can access electronic materials using software called screenreaders. Screenreaders convert the text in electronic teaching materials into speech so that blind and partially sighted students can listen to the material. However, screenreaders cannot provide any information about images, and electronic teaching materials contain many images. The only way that a screenreader can convey the information in an image is having someone create a description of the image which can then be read out.

Our research is about creating an online crowdsourcing system that will allow sighted students to very easily add descriptions to images on the electronic teaching materials for their courses. Students would be able to describe as many or as few images as they wanted to. The system will include instruction on how to create good image descriptions.

The University of York is trying to be particularly welcoming to blind and partially sighted students, and the History Department has an excellent record of accepting blind and partially sighted student, so we thought plotting this idea in History would be a good idea.

At the moment we are interested to find out what would motivate sighted students to participate in such a system. The questionnaire asks a simple set of questions about your thoughts about the system, you should find it interesting and it should take no longer than 10 minutes to complete. All people who complete the questionnaire by the end of term will be entered in a prize draw for 10 £10 Amazon gift vouchers.

C.3.2. Phase 2: Recruitment email

Dear Students,

We are piloting the "DescribeIT" project; an online system to support blind and partially sighted students, by providing access to images that are part of the PowerPoint slides that lecturers use in teaching. *More information can be found in the project Info page*

The online project will allow sighted students to very easily add descriptions to images on the electronic teaching materials of "Using Visual Materials in Historical Research" course. Students would be able to describe as many or as few images as they wanted to. The system will include instruction on how to create good image descriptions.

The University of York is trying to be particularly welcoming to blind and partially sighted students, and the History Department has an excellent record of accepting blind and partially sighted student, so we thought running this project with the History students would be a good idea.

I would be very grateful if you would be willing to take part in our project. If you are interested, please use the link below to access the project.

[****The Project Link*****]

Many thanks

Helen

Helen Petrie Phd AFBPsS CPsychol FRSA

C.3.3. Round 2: Phase 2: Recruitment email

Dear Students,

We are contacting you because you had previously showed interest in supporting your blind and partially sighted peers. We would like to invite you to take part in the "DescribeIT" project; an online system to support blind and partially sighted students, by providing access to images that are part of the PowerPoint slides that lecturers use in teaching. More information can be found in the project Info page

The online project will allow you to very easily add descriptions to images on the electronic teaching materials of your last year course "Using Visual Materials in Historical Research". You would be able to describe as many or as few images as you want. The system will include instruction on how to create good image descriptions.

The University of York is trying to be particularly welcoming to blind and partially sighted students, and the History Department has an excellent record of accepting blind and partially sighted student, so we thought running this project with the History students would be a good idea.

I would be very grateful if you would be willing to take part in our project. If you are interested, please use the link below to access the project.

[****The Project Link*****]

Many thanks

Helen

Helen Petrie Phd AFBPsS CPsychol FRSA

C.3.4. Examples of Students' Comments:

Table C.25 Comments from students about Altruism

Likert Ratings	Students' comments
2	I don't feel I'm obligated to participate (P12)
3	While Ii do enjoy helping others, I do this enough for those that I know and care about. I don't really have the time to help even more. (P15)
4	I just think its a practical way to help someone. (P26)
5	I think there would be genuine sense of helping other students if the project was launched. (P33)
6	I like the satisfaction of helping others (P23)
7	Helping people is very important and the ultimate goal really! (P27)

Table C.26 Comments from students about improving skills

Likert Ratings	Students' comments
1	this project wouldn't improve my skills (P12)
2	This was something I hardly considered. (P6)
3	I'm not sure it would aid my academic skills (P13)
4	Anything that involves writing is helpful to my degree but again I have little time to spare (P34)
5	It would probably make me look more closely at sources. (P26)
6	Picture sources are often easy to overlook in particular details, by needing to describe them I feel this could improve my own analysis as well as help others. (P16)
7	I'm very competitive and dislike feeling like there's a level of skill that I haven't reached. (P24)

Table C.27 Comments from students about money

Likert Ratings	Students' comments
1	Because the main aim of the project is to help others (P36)
2	The Amazon Voucher raffle is a nice touch but it is certainly not the be all and end all. (P6)
3	Money is nice but I would feel a bit of a cheat taking money from a charity as long as it was a doable amount of time I was working for them. (P27)
4	Whilst it would be nice to be paid this would not be the main motivation for participating. (P16)
5	It would provide a motive for anyone to help. (P2)
6	This can be a good way of bringing attention to the project and getting people 'through the door' (P14)
7	Because being a student is expensive (P1)

Table C.28 Comments from students about enhancing job opportunities

Likert Ratings	Students' comments
1	I have hardly considered this. (P6)
2	Don't want to go into a career relating to this field (P4)
4	While this would be useful and desired, I currently can't see how this would enhance my career prospects. (P15)
6	It would demonstrate the willingness to help others and also improve my communication skills. (P2)
7	Because anything that will help me get a job is worth taking part in (P1)

Table C.29 Comments from students about being connected to others

Likert Ratings	Students' comments
2	Don't understand how it would connect me to other students on my course, unless you took part together in the same place etc. (P4)
4	Already know a lot of people at university but meeting new people is always nice. (P27)
6	There are many of us here and being connected with others would enhance my degree experience (P1)

Table C.30 Comments from students about fun and entertainment

Likert Ratings	Students' comments
1	I think it would get boring after a real descriptions. However I would still do it as it's important to other students. (P4)
2	Although I enjoy my degree, it doesn't mean that I want to do it at all times; I need to have a break and this would not aid in that. (P15)
4	It could be fun, I'm not overly bothered if it's not though. (P20)
6	Seems interesting (P28)
7	Because studying is hard and it is nice to have a break (P1)

Table C.31 Comments from students about passing the time

Likert Ratings	Students' comments
1	I don't have any spare time. (P4)
2	I've got better things to do (P12)
6	It would provide a productive break from work. (P2)

Table C.32 Comments from students about contributing to a large project

Likert Ratings	Students' comments
1	I have a lot of other commitments. (P26)
3	It would be a nice thing to be a part of but merely being a 'part' doesn't mean much in today's society. (P15)
4	Being part of something big is rewarding. (P35)
5	It would be good to be part of something that will make such a difference (P4)
6	Knowing I am helping other with this project would make me feel good (P1)
7	It always makes you feel fulfilled to be part of a larger accomplishment. (P20)

Table C.33 Comments from students about the social recognition

Likert Ratings	Students' comments
1	I'm not particularly bothered about the recognition. (P20)
2	Social recognition isn't that important to me; I have greater things to worry about. (P15)
4	Social recognition isn't very important to me (P1)

Table C.34 Comments from students about drawing attention to skills

Likert Ratings	Students' comments
1	I'm not a show off (P4)
2	no skills required (P12)
3	I believe I would more interested in the improvement of my skills than the recognition of them (P16)
5	Would help me develop my own skills (P1)

Table C.35 Comments from students about being in competition

Likert Ratings	Students' comments
1	I don't feel I have to compete with other students (P4)
2	I am not very competitive (P16)
3	I'm very competitive but it's not of particular importance in this task - at least I don't think so. (P20)
4	I'm not really bothered about being in competition with someone else (P1)

Table C.36 Comments from students about getting academic credits

Likert Ratings	Students' comments
1	I wouldn't be able to for this project (P12)
3	It would be excellent if this was at all possible. (P20)
4	I'd be intrigued as to how this would work and might be motivated, depending on how it worked. (P15)
5	Whilst welcomed and a form of encouragement, the credits would not be the primary reason for joining. (P16)
6	That would definitely make it worth the time and effort, as it would distract from uni work but gaining credits makes it worth it (P4)
7	Because I would like to get the highest grade possible (P1)

Appendix D
Chapter 7

D.1. Study 6

D.1.1. Recruitment email

Dear Student,

We are piloting the "DescribeIT" project; an online system to support blind and partially sighted students, by providing access to images that are part of the PowerPoint slides that lecturers use in teaching. More information can be found in the project info page.

The online project will allow sighted people like you to very easily add descriptions to images in the digital teaching materials. No specific knowledge is required for you to be able to describe the images. You will be able to describe as many or as few images as you wanted to. The system will include instructions on how to create good image descriptions.

You will be paid 20p per description, payable as an Amazon gift voucher. A total of 27 images are available for descriptions in the pilot project, meaning that you can earn up to £5.40. You could increase the amount you earn by also participating in an online questionnaire about describing images for blind and partially sighted students. It takes about 5 minutes to complete, and everyone who completes the questionnaire will be entered into a prize draw for one of 10 Amazon gift vouchers, each valued at £10. All information in the questionnaire will be completely confidential and anonymous.

We would be very grateful if you would be willing to take part in this project. If you are interested, please use the link below to access the project:

[***Project link ****]

If you have any questions, please contact Fatma Layas at fal503@york.ac.uk.

Thank you for taking the time to consider our request.

Fatma Layas, PhD student

and Helen Petrie Phd AFBPsS CPsychol FRSA

Appendix E
Chapter 8

E.1. Study 8

E.1.1. Recruitment email

Dear Student,

One of my PhD students, Fatma Layas, is piloting an application called "TagIT", which is about improving access to digital teaching material for blind and partially sighted students. We are asking you to help us in a small way with this pilot.

One big problem for blind and partially sighted students is that lecturers use many images in their teaching materials, and if you cannot see them (clearly) you may miss out on important information for learning. The TagIt application allows sighted students to tag the objects in the images in digital teach materials very easily which can then be built up into descriptions of the images. More information about the importance of this application for blind and partially sighted students can be found in the project information page, along with tips to help you tag images and an example of a set of tags for an image.

The TagIt application will allow you to tag as many or as few images as you wish. Participating in this pilot will also improve your skills in analysing and tagging images which will make your blog articles, YouTube videos and images easy to find by search engines. Following the image tagging tips will help you learn how to provide good image tags.

Please try it out, you can access it via this link:

[***Project link ****]

If you have any questions, please do not hesitate to contact Fatma Layas at Fal503@york.ac.uk.

Thank you in advance!

Helen Petrie Phd AFBPsS CPsychol FRSA

E.1.2. Tips To Help You Tag Images:

What is the content of the image: Create simple descriptions by putting what you can see in the image. This can be done by breaking the image down into its basic components. Bear in mind that your tags will be used by students to create more detailed descriptions so you do not have to be overly specific. You can use the information provided with each image to create tags.

Keep your tags relevant to the image content: Keep your tags directly related to the image content. Do not throw in random keywords as this can be considered as tag spamming.

Avoid words with multiple meanings: words with multiple meanings can be confusing, so if you use one extend your tag to add the right meaning (e.g. the word “Bank” it can mean the land alongside or sloping down to a river (“willows lined the bank”) or lake or a slope, mass, or mound of a particular substance (“a bank of clouds”).

Multiple tags: you can add as much tags as you wish as long as they are relevant to the content of the image.

Appendix F
Chapter 9

F.1. Study 10

F.1.1. Online Questionnaires Pre-Study Phase:

1. The information and demografic sections the same as in B.1.3
2. Difficulty level:

What do you think will be the difficulty level of the image description task?

Extremely Difficult ... Extremely Easy

F.1.2. Online Questionnaires Pre-Training Phase:

1. Difficulty level:

Now you tried the image description task, what do you think was the difficulty level of it?

Extremely Difficult ... Extremely Easy

*

2. Situational Motivation Scale (SIMS), the same as in C.1.3

F.1.3. Online Questionnaires Post-Training Phase:

1. Difficulty level:

Now you tried the image description task after the training session, what do you think was the difficulty level of it?

Extremely Difficult ... Extremely Easy

*

2. Situational Motivation Scale (SIMS), Interaction with Disabled Persons Scale (IDP), Self-Report Altruism Scale (SRA) the same as in C.1.3

F.1.4. Information document the same as in C.2.5

F.1.5. The project information page

DescribeIT: supporting blind and partially sighted students by describing images

Help us test DescribeIT, a crowdsourcing project to improve access to online teaching materials for blind and partially sighted students.

Online teaching materials contain many images. Some of these images are only decoration. However, many images are important to understanding the material and being able to learn it. Blind and partially sighted students are not be able to learn effectively if they do not know what is in the images, but lecturers do not have time to provide an explicit explanation of every image they use in their teaching materials.

The DescribeIT application allows participants to provide descriptions of images for blind and partially sighted students.

**** The quality control instructions ****

Guidelines for Describing Images for Blind and Partially Sighted Students

The same as in A.1

References

Ajjan, H. & Hartshorne, R. (2008). Investigating Faculty Decisions to Adopt Web 2.0 Technologies: Theory And Empirical Tests. *The internet and higher education, 11(2), 71-80.*

Anderson, J. & Clark, J. (2015). Effectiveness of Online Altruistic Viral Videos on Altruistic Attitudes and Behaviors.

Antin, J. & Shaw, A. (2012). Social Desirability Bias and Self-Reports of Motivation: A Study of Amazon Mechanical Turk in the US and India. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)* (May 5-10, Austin, TX, USA). 2925-2934. New York: ACM Press.

Armitage, C. J. & Conner, M. (2001). Efficacy of The Theory of Planned Behaviour: A Meta-Analytic Review. *British journal of social psychology*, 40(4), 471-499.

Arnocky, S., Piché, T., Albert, G., Ouellette, D., & Barclay, P. (2017). Altruism Predicts Mating Success in Humans. *British Journal of Psychology*, 108(2), 416-435.

Bigham, J. P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R. C., ... & Yeh, T. (2010, October). Vizwiz: Nearly Real-Time Answers to Visual Questions. *In Proceedings of the 23nd annual ACM symposium on User interface software and technology* (pp. 333-342). ACM.

Bigham, J.P., Ladner, R. E. & Borodin, Y. (2011). The Design of Human-Powered Access Technology. *In Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11)* (October 24 – 26, Dundee Scotland). 3-10. New York: ACM Press.

Bowser, A., Hansen, D., He, Y., Boston, C., Reid, M., Gunnell, L., & Preece, J. (2013, October). Using Gamification to Inspire New Citizen Science Volunteers. *In Proceedings of the first international conference on gameful design, research, and applications* (pp. 18-25). ACM.

Brabham, D. C. (2008). Moving The Crowd At Istockphoto: The Composition Of The Crowd And Motivations For Participation In A Crowdsourcing Application. *First Monday* 13(6).

Brabham, D. C. (2010). Moving The Crowd At Threadless: Motivations For Participation In A Crowdsourcing Application. *Information, Communication & Society* 13(8): 1122-1145.

Brabham, D. C. (2012). Motivations for Participation in A Crowdsourcing Application to Improve Public Engagement in Transit Planning. *Journal of Applied Communication Research* 40(3): 307-328.

Brooke, J. (1996). SUS a Quick and Dirty Usability Scale. *Usability evaluation in industry*, 189(194), 4-7.

Brooke, J. (2013). SUS: a Retrospective. *Journal of usability studies*, 8(2), 29-40.

Bussell, H. and Forbes, D. (2002) Understanding The Volunteer Market: The What, Where, Who and Why of Volunteering, *International Journal of Nonprofit and Voluntary Sector Marketing,* 7 (3), pp.244-257.

Cardonha, C.; Gallo, D.; Avegliano, P.; Herrmann, R.; Koch, F.; and Borger, S. 2013. A Crowdsourcing Platform for the Construction of Accessibility Maps. *In Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*, 26-29. New York: ACM Press.

Chandler, D. & Horton, J. J. (2011). Labor Allocation in Paid Crowdsourcing: Experimental Evidence On Positioning, Nudges And Prices. *Human Computation*, 11, 11.

Chandler, D. & Kapelner, A. (2013). Breaking Monotony with Meaning: Motivation In Crowdsourcing Markets*. Journal of Economic Behavior & Organization*, 90, 123-133.

Chen, Y. (2013). Crowd Sourcing Descriptions of Images on The Web for Visually Impaired People. Department of Computer Science, *The university of York. Master in Human-Centred Interactive Technologies:* 162.

Clary, E. G. & Snyder, M. (1991). A Functional Analysis of Altruism and Prosocial Behavior: The Case of Volunteerism. In M. S. Clark (Ed.), *Review of personality and social psychology, Vol. 12. Prosocial behavior* (pp. 119-148). Thousand Oaks, CA: Sage Publications.

Clary, E. G., Snyder, M., & Stukas, A. A. (1996). Volunteers' Motivations: Findings from a National Survey. *Nonprofit and voluntary sector quarterly*, *25*(4), 485-505.

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., & Miene, P. (1998). Understanding and Assessing The Motivations Of Volunteers: A Functional Approach. *Journal of personality and social psychology*, *74*(6), 1516.

Deci, E. L. & Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. New York: Plenum.

Deci, E. & Ryan, R. M. (2012). Self-Determination Theory. *Handbook of theories of social psychology*, Los Angeles, SAGE. ISBN: 9780857029607

Deterding, S., Khaled, R., Nacke, L., and Dixon, D., (2011), Gamification: Toward a Definition, *CHI*, Vancouver, BC, Canada. ACM 978-1-4503-0268-5/11/05.

Difallah, D. E., Catasta, M., Demartini, G., Ipeirotis, P. G., & Cudré-Mauroux, P. (2015, May). The Dynamics of Micro-Task Crowdsourcing: The Case of Amazon MTurk. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 238-247). International World Wide Web Conferences Steering Committee.

Donaldson, S. I. & Grant-Vallone, E. J. (2002). Understanding Self-Report Bias in Organizational Behavior Research. *Journal of business and Psychology*, *17*(2), 245-260.

Dow, S., Kulkarni, A., Klemmer, S., & Hartmann, B. (2012, February). Shepherding the Crowd Yields Better Work. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1013-1022). ACM.

Dumitrache, A., Aroyo, L., Welty, C., Sips, R. J., & Levas, A. (2013, October). Detective: Combining Gamication Techniques and Crowdsourcing to Create a Gold Standard in Medical Text. In *Proceedings of the 1st International Conference on Crowdsourcing the Semantic Web-Volume 1030* (pp. 16-31). CEUR-WS. org.

Estellés-Arolas, E. & González-Ladrón-De-Guevara, F. (2012). Towards an Integrated Crowdsourcing Definition. *Journal of Information science*, 38(2), 189-200.

Faullant, R., Holzmann, P., & Schwarz, E. J. (2016). Everybody Is Invited but Not Everybody Will Come The Influence Of Personality Dispositions on Users' Entry Decisions For Crowdsourcing Competitions. *International Journal of Innovation Management*, *20* (06), 1650044.

Feyisetan, O., Simperl, E., Van Kleek, M., & Shadbolt, N. (2015, May). Improving Paid Microtasks Through Gamification and Adaptive Furtherance Incentives. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 333-343). ACM.

Fishbach, A., & Zhang, Y. (2008). Together or Apart: When Goals and Temptations Complement Versus Compete. *Journal of personality and social psychology*, *94*(4), 547.

Fisher, R. J. (1993). Social Desirability Bias and The Validity of Indirect Questioning. *Journal of consumer research*, 20(2), 303-315.

Fogg, B. J. (2009, April). A Behavior Model for Persuasive Design. In *Proceedings of the 4th international Conference on Persuasive Technology* (p. 40). ACM.

Geen, R. G. (1995). Human Motivation: A Social Psychological Approach. *Thomson Brooks/Cole Publishing Co*. Belmont, CA, US.

George, J. F. (2004). The Theory of Planned Behavior And Internet Purchasing. *Internet research*, *14*(3), 198-212.

Gething, L. (1994). The Interaction with Disabled Persons Scale. *Journal of Social Behavior and Personality*, 9(5), 23.

Goncalves, J., Ferreira, D., Hosio, S., Liu, Y., Rogstadius, J., Kukka, H., & Kostakos, V. (2013, September). Crowdsourcing on The Spot: Altruistic Use of Public Displays, Feasibility, Performance, and Behaviours. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (pp. 753-762). ACM.

Goncalves, J., Hosio, S., Rogstadius, J., Karapanos, E., & Kostakos, V. (2015). Motivating Participation and Improving Quality of Contribution in Ubiquitous Crowdsourcing. *Computer Networks*, *90*, 34-48.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and emotion*, 24(3), 175-213.

Hackman, J. R., & Oldham, G. R. (1980). Work Redesign, Addison-Wesley, Reading Mass.

Hara, K.; Le, V. and Froehlich, J. (2012). A Feasibility Study of Crowdsourcing and Google Street View to Determine Sidewalk Accessibility. *In Proceedings of the 14th international ACM SIGACCESS conference on Computers and Accessibility*, 273-274. New York: ACM Press

Hara, K;, Azenkot, S.; Campbell, M.; Bennett, C. L.; Le, V.; Pan-nella, S. and Froehlich, J. E. (2013). Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View. *In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, 6-13.New York: ACM Press.

Harter, S. (1981). A New Self-Report Scale of Intrinsic Versus Extrinsic Orientation in The Classroom: Motivational and Informational Components. *Developmental psychology*, *17*(3), 300.

Higher Education Statistical Agency, (2013), available at: https://www.hesa.ac.uk/ last access (01 June 2017)

Hofstede, G. 1983. The Cultural Relativity of Organizational Practices and Theories. *Journal of International Business Studies*, Fall, 75–90.

Hofstede, G. 2001. Culture's Consequences: *Comparing Values, Behaviors, Institutions, and Organizations across Nations (2nd ed.).* Thousand Oaks, CA; London: Sage Publications.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14(3), 399-425.

Hossain, M. (2012). Crowdsourcing: Activities, Incentives and Users' Motivations To Participate. *International Conference on Innovation, Management and Technology Research (ICIMTR2012), (May 21-22, Malacca, Malaysia)* 501-506. IEEE.

Houle, B. J., Sagarin, B. J., & Kaplan, M. F. (2005). A Functional Approach to Volunteerism: Do Volunteer Motives Predict Task Preference? *Basic and applied social psychology*, *27*(4), 337-344.

Howe J. 2006, Crowdsourcing: a definition. Available at: http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html. Last access (01 June 2017)

Hsieh, H. F. & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative health research*, *15*(9), 1277-1288.

Ipeirotis, Panagiotis G., Demographics of Mechanical Turk (March 2010). NYU Working Paper No. CEDER-10-01. Available at SSRN: http://ssrn.com/abstract=1585030

John, O. P., & Srivastava, S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. *Handbook of personality: Theory and research*, *2* (1999), 102-138.

Jurgens, D., & Navigli, R. (2014). It's All Fun and Games Until Someone Annotates: Video Games With A Purpose For Linguistic Annotation. *Transactions of the Association for Computational Linguistics*, *2*, 449-464.

Kaufmann, N., & Schulze, T. (2011). Worker Motivation in Crowdsourcing And Human Computation. *Education*, *17*(2009), 6.

Kaufmann, N., Schulze, T. & Veit, D. (2011). More Than Fun and Money: Worker Motivation In Crowdsourcing: A Study On Mechanical Turk. *In Proceedings of the Seventeenth Americas Conference on Information Systems (August 4-7 2011, Detroit, Michigan)*. 1-11. Association for Information Systems (AMCIS).

King, M. F., & Bruner, G. C. (2000). Social Desirability Bias: A Neglected Aspect of Validity Testing. *Psychology and Marketing*, *17*(2), 79-103.

Krull, D. S., et al. (1999). The Fundamental Fundamental Attribution Error: Correspondence Bias In Individualist And Collectivist Cultures. *Personality and Social Psychology Bulletin* 25(10): 1208-1219.

Kuznetsov, S. (2006). Motivations of Contributors to Wikipedia. *ACM SIGCAS Computers and Society* 36(2): 1.

Lakhani, K. R. (2008). InnoCentive. com (A). *Harvard Business School Case*, (608-170)

Lakhani, K. R., and Wolf, R. G. (2005) Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects, *In Perspectives on free and open source software*, MIT Press.

Lakhani, K. R., Jeppesen, L. B., Lohse, P. A. & Panetta, J. A (2007). The Value of Openess In Scientific Problem Solving. *Division of Research, Harvard Business School*.

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). Research Methods in Human-Computer Interaction. *Morgan Kaufmann*.

Lee, T. Y., Dugan, C., Geyer, W., Ratchford, T., Rasmussen, J. C., Shami, N. S., & Lupushor, S. (2013, July). Experiments on Motivational Feedback for Crowdsourced Workers. In *ICWSM*.

MacLean, D. and Gannon, P. M. 1995. Measuring Attitudes Toward Disability: The Interaction with Disabled Persons Scale Revisited. *Journal of Social Behavior and Personality*, 10(4), 791

Mao, A., Kamar, E., & Horvitz, E. (2013, November). Why Stop Now? Predicting Worker Engagement in Online Crowdsourcing. In *First AAAI Conference on Human Computation and Crowdsourcing*.

Martin, D., Hanrahan, B. V., O'Neill, J., & Gupta, N. (2014, February). Being a Turker. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 224-235). ACM.

Mason, W., & Suri, S. (2012). Conducting Behavioral Research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.

McCrae, R. R., & John, O. P. (1992). An Introduction to The Five-Factor Model and Its Applications. *Journal of personality*, *60* (2), 175-215.

McDonald, J. D. (2008). Measuring Personality Constructs: The Advantages and Disadvantages of Self-Reports, Informant Reports and Behavioural Assessments. *Enquire*, *1*(1), 1-18.

Morschheuser, B., Hamari, J., & Koivisto, J. (2016, January). Gamification in Crowdsourcing: A Review. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on* (pp. 4375-4384). IEEE.

Nov, O. (2007). What Motivates Wikipedians? *Communications of the ACM*, 50(11), 60 – 64.

Oreg, S. & O. Nov (2008). Exploring Motivations for Contributing to Open Source Initiatives: The Roles of Contribution Context and Personal Values. *Computers in human behavior* 24(5): 2055-2073.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running Experiments on Amazon Mechanical Turk. *Judgment and Decision making*, 5(5), 411-419.

papworthtrust.org.uk, (2016) Disability in the United Kingdom 2016 Facts and figures, available at: http://www.papworthtrust.org.uk/sites/default/files/Disability%20Facts%20and%20Figures%202016.pdf [accessed 21 June 2017]

Pelling, E. L. & White, K. M. (2009). The Theory of Planned Behavior Applied to Young People's Use Of Social Networking Web Sites. *CyberPsychology & Behavior*, *12*(6), 755-759.

Petrie, H., Harrison, C. & Dev, S. 2005. Describing Images on The Web: A Survey of Current Practice and Prospects For The Future. *Proceedings of 3rd International Conference on Universal Access in Human-Computer Interaction. (July 22 – 27, Las Vegas)*.

Pritchard, R. & Ashwood, E. (2008). *Managing Motivation: A Manager's Guide to Diagnosing and Improving Motivation.* New York: Routledge Academic

Raskind, M., Hoffman, B., Hartley, K., & Boone, R. (2005). Reaching Accessibility Guidelines for Creating and Refining Digital Learning Materials. *Intervention in School and Clinic*, 40(3), 171-176.

Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., & Vukovic, M. (2011). An Assessment of Intrinsic And Extrinsic Motivation on Task Performance In Crowdsourcing Markets. *ICWSM*, *11*, 17-21.

Rosenthal, R. & Rosnow, R. L. (2009). Artifacts in Behavioral Research. *Oxford University Press*.

Ross, J., Irani, L., Silberman, M., Zaldivar, A., & Tomlinson, B. (2010, April). Who Are The Crowdworkers?: Shifting Demographics In Mechanical Turk. In *CHI'10 extended abstracts on Human factors in computing systems* (pp. 2863-2872). ACM.

Ross, L. & Nisbett, R.E. (2011). The Person and The Situation: Perspectives on Social Psychology. *London: Pinter and Martin.*

Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The Altruistic Personality and The Self-Report Altruism Scale. *Personality and individual differences*, *2*(4), 293-302.

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: a Database And Web-Based Tool For Image Annotation. *International journal of computer vision*, *77*(1), 157-173.

Ryan, R. M. (Ed.). (2012). The Oxford Handbook of Human Motivation. Oxford University Press.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and The Facilitation Of Intrinsic Motivation, Social Development, and Well-Being. *American psychologist*, *55*(1), 68.

Rynes, S. L., Gerhart, B., & Minette, K. A. (2004). The Importance of Pay In Employee Motivation: Discrepancies Between What People Say And What They Do. *Human resource management*, *43*(4), 381-394.

Sauro, J. (2011). Measuring Usability with The System Usability Scale (SUS). Available at: https://measuringu.com/sus/ [accessed 29 May 2017]

Schwarz, N. (1999). Self-Reports: How The Questions Shape The Answers. *American psychologist*, *54*(2), 93.

Shaw, A. D., Horton, J. J., & Chen, D. L. (2011, March). Designing Incentives for Inexpert Human Raters. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 275-284). ACM.

Singer, A. (2013). Giving Practices in Islamic Societies. *social research*, *80*(2), 341-358.

Smith, D., et al. (2013). "How Can Entrepreneurs Motivate Crowdsourcing Participants?" Technology Innovation Management Review(February: Platforms, Communities, and Business Ecosystems).

Smith, W. T., Catney, C. M., Rickles, N. M., Hermansen-Kobulnicky, C. J., Broeseker, A. E., Garvan, C. W., & Kimberlin, C. L. (2014). Pharmacy student comfort in communicating with persons with disabilities. *Currents in Pharmacy Teaching and Learning*, *6*(2), 175-184.

Snodgrass, J. G., Levy-Berger, G., & Haydon, M. (1985). *Human experimental psychology* (Vol. 395). New York: Oxford University Press.

Splendiani, B., & Ribera, M. (2014). Accessible Images in Computer Science Journals. *Procedia Computer Science*, 27, 9-18.

Splendiani, B., Ribera, M., & Centelles Velilla, M. (2014). Are Figures Accessible In Mathematics Academic Journals*?. In Comunicació presentada a: 29th Annual International Technology and Persons with Disabilities Conference. The 2014 CSUN Conference was held in San Diego*, CA from March 17 to March 22, 2014.

Takagi, H., Kawanaka, S., Kobayashi, M., Itoh, T. & Asakawa, C. (2008). Social Accessibility: Achieving Accessibility Through Collaborative Metadata Authoring*. In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '08) (October 13–15, 2008, Halifax, Nova Scotia, Canada). 193-200*. New York: ACM Press.

Thaler, S., Simperl, E., & Wölger, S. (2012). An Experiment in Comparing Human-Computation Techniques. *IEEE Internet Computing*, *16*(5), 52-58.

Touré-Tillery, M., & Fishbach, A. (2014). How to Measure Motivation: A Guide For The Experimental Social Psychologist. *Social and Personality Psychology Compass*, *8*(7), 328-341.

Vallerand, R. J. (1997). Toward A Hierarchical Model of Intrinsic and Extrinsic Motivation. *Advances in experimental social psychology*, *29*, 271-360.

Von Ahn, L. & L. Dabbish (2004). Labeling Images with a Computer Game. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04) (April 24–29, 2004, Vienna, Austria).* 193-200. New York: ACM Press.

Von Ahn, L., Liu, R., & Blum, M. (2006, April). Peekaboom: A Game for Locating Objects in Images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 55-64). ACM.

Wald, M. (2010). Synote: Accessible and Assistive Technology Enhancing Learning For All Students. *Computers Helping People with Special Needs*, Springer: 177-184.

Weiss, A. (1995) Human Capital vs. Signalling Explanations of Wages, *Journal of Economic Perspectives*, 9, 4, 133-54

Yuen, M. C., King, I., & Leung, K. S. (2012, August). Task Recommendation in Crowdsourcing Systems. In *Proceedings of the first international workshop on crowdsourcing and data mining* (pp. 22-26). ACM.

Zhang, P. (2007). Toward A Positive Design Theory: Principles for Designing Motivating Information And Communication Technology. In *Designing information and organizations with a positive lens* (pp. 45-74). Emerald Group Publishing Limited.

Zheng, H., et al. (2011). Task Design, Motivation, And Participation In Crowdsourcing Contests. *International Journal of Electronic Commerce 15(4): 57-88.*