

1st Supervisor: Dr. Ariane von Raesfeld Meijer
Universiteit Twente
Chair of Business Development in a Network Perspective

2nd Supervisor: Prof. Dr. Knut Blind
Technische Universität Berlin
Chair of Innovation Economics

Master Thesis

The influence of social identity on creative crowdsourcing engagement

Starke, Ina
Meininger Straße 9
10823 Berlin
inastarke@gmail.com
Matriculation numbers: s1360871 (Twente) / 338464 (Berlin)
Innovation Management and Entrepreneurship

Abstract

The objective of this thesis is to empirically test the proposition by Piyathasanan et al. (2011) stating that a high social identity (self categorization, self-esteem, affective commitment) positively influences crowdsourcing engagement. This proposition has not been empirically tested in the literature, yet. An online questionnaire was sent to people in Estonia and Switzerland to fill this research gap. The method of a regression analysis was applied to evaluate the questionnaire. The empirical results show that self-esteem is the only factor of social identity which significantly influences people's crowdsourcing engagement. Consequently, it would be wise of organizations to make their members feel comfortable and with that increase their self-esteem in order to be rewarded for that in terms of an increased willingness from the side of members to act in favor of their organization.

Table of Content

1. Relevance of crowdsourcing	2
2. Introduction of the research hypothesis	4
3. Literature review.....	7
3.1. Crowdsourcing	7
3.2. Social identity and its relationship to crowdsourcing engagement.....	10
4. Method	23
4.1. Research setting.....	23
4.2. Description of the sample	24
4.3. Operationalization of variables.....	25
4.3.1. Processing of variables.....	25
4.3.2. Measuring social identity, crowdsourcing engagement and control variables	26
4.4. Steps for data analysis	37
5. Results.....	38
5.1. Correlation analysis.....	38
5.2. Cook's distance and Mahalanobis distance	43
5.3. Four regression models.....	45
6. Conclusion and discussion	50
Literature	52
Appendix	55
SPSS syntax.....	55
Assumptions for a regression analysis.....	64
Online questionnaire	73

Table of Figures

Figure 1: Conceptual model of Piyathasanan et al. (2011) adopted from: (Piyathasanan et al., 2011, p. 1)	5
Figure 2: Sample characteristics - summarizing table	24
Figure 3: Scale for self categorization - partly adapted from (Fred Mael & Tetrick, 1992, p. 817)	27
Figure 4: Scale for group self-esteem - partly adapted from (Ellemers et al., 1999, p. 379)	28
Figure 5: Scale for affective commitment - partly adapted from (Allen & Meyer, 1990, p. 6)	29
Figure 6: Reliabilities of the scales – self categorization, self-esteem, affective commitment, social identity	31
Figure 7: Item for creative crowdsourcing engagement	31
Figure 8: Control variables – personal information	33
Figure 9: Control variables - specifically relevant for this study	34
Figure 10: Correlation table	42
Figure 11: Cook's and Mahalanobis distance - regression models I-IV	44
Figure 12: Regression analyses - regression models I-IV	46
Figure 13: Summary of the results from the regression analyses	50

1. Relevance of crowdsourcing

Nowadays companies are constantly under pressure not only to come up frequently with innovative ideas, but also to increase their overall performance to stay competitive on the market. Trying to meet these requirements many companies face the challenge of scarce resources in terms of workforces. Thanks to technical advancements through the internet crowdsourcing¹ - a term and phenomenon which became popular since a few years ago - could be a way to handle these problems. Following the definition of crowdsourcing by Howe (2006) the idea is to open up the inside world of the firm with all of its problems, tasks and challenges to the outside world (Howe, 2006a). This idea is closely connected to the open innovation paradigm coined by Chesbrough (2003) which says that innovative ideas for the market do not necessarily only come from inside the company, but can also be produced externally (Chesbrough, 2003, p. 43). Thereby, companies are able to give problems concerning persisting concepts and ideas for new solutions away to people who are originally unknown and maybe even unconnected to the respective firm. This act implies the opportunity for companies to find the one person or a group of persons among the many within the crowd who has a solution for the problem the company could not solve with its internal resources. The theoretical assumption behind this is that individuals within a heterogeneous crowd are able to develop ideas and solutions which equal or are even better than the quality of internal company experts (Michelis & Schildhauer, 2012, p. 136). Surowiecki (2005) calls this “the wisdom of the crowds” and explains in his book that the crowd has the potential to outperform the smartest people when it comes to problem solving or answering specific questions (Surowiecki, 2005, p. 11). Brabham (2013) further enumerates the conditions proposed by Surowiecki (2005) under which a crowd can outperform experts: Independence of individuals in the group, diversity of the group and the aggregation of individual outputs (Brabham, 2013). Kristensson et al. (2004) found that ordinary users have the potential to develop more original and valuable ideas compared to internal professionals (Kristensson et al., 2004, p. 11). Additionally, also Poetz and Schreier (2012) provide empirical evidence for the fact that user ideas score higher for customer benefit and novelty (Poetz & Schreier, 2012, pp. 17-18; 20). Having this background information in mind crowdsourcing seems to turn the traditional understanding of the

¹ Definition of “crowdsourcing”: “Simply defined, crowdsourcing represents 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, 2006a) This definition will be explained in detail later in this thesis.

relationship between firms and its employees upside-down. Developing new ideas and solutions is no longer restrained to internal company resources but can also be achieved by openly consulting an external crowd of people who show interest in the problem at hand. This prospect opens up completely new opportunities especially for smaller and medium sized companies with constrained internal resources. Consequently, it would be interesting to know what motivates individuals in the crowd to contribute to crowdsourcing projects. A basic understanding of the motivational factors to participate in crowdsourcing would enable companies to directly address the potential crowd or individuals within it. Therefore, academics as well as practitioners asked themselves: What motivates people to contribute something to crowdsourcing projects which are initiated by a company they might not know and for which they sometimes might not even get paid?

In the following lines a short overview of the structure of this thesis is given: Firstly, my hypothesis will be introduced with the help of the article by Piyathasanan et al. (2011) which serves as a basis for this master thesis. Secondly, this section is followed by the literature review which will contain definitions of the terms crowdsourcing engagement and social identity together with its three sub-factors self categorization, group self-esteem and affective commitment. Additionally, different sources will be used in order to explore the influence of social identity and its three sub-factors on people's behavior towards the group they derive their social identity from. It will be shown that social identity together with its three sub-factors positively influences people's engagement for the groups they belong to. On the basis of this finding and the fact that none of the authors so far explicitly tested the influence of social identity and its three sub-factors on crowdsourcing engagement the research hypothesis will be presented and justified. Thirdly, in the method part several points will be discussed: By starting with the research setting it will be explained why an online questionnaire was chosen to conduct the study and where this questionnaire was spread. Additionally, my research sample will be described by listing its characteristics. In connection with the sample description also the assumptions of normal distribution and homogeneity of variance will be tested. Furthermore, the operationalization of all variables included in this study will be explained. Moreover, the process of developing and measuring the variables and scales for the constructs social identity with its three sub-factors and crowdsourcing engagement will be presented. In this context the resulting scales for self categorization, self-esteem, affective commitment and social identity will be tested for their respective reliabilities. Here, also the control variables used in this study will be introduced. As a last paragraph of the method part steps which are need to be taken in order to evaluate the data will be enumerated. Fourthly, the several steps for the development and presentation of the

results are needed: Conducting a correlation analysis, testing for Cook's and Mahalanobi's distance, checking all of the assumptions for pursuing a regression analysis and performing four regression analyses². Last but not least, the results of the analyses will be summarized, conclusions will be drawn and managerial implications will be given.

2. Introduction of the research hypothesis

The question what motivates people to contribute to crowdsourcing was also of interest for Piyathasan et al. (2011) when they published their paper "Social Identity and Motivation for Creative Crowdsourcing and Their Influence on Value Creation for the Firm". Within their paper the authors proposed a conceptual model which according to their own words needs to be tested empirically (Piyathasan et al., 2011, p. 5). After having reviewed the literature it can be concluded that until today no one explicitly tested the first proposition as stated by Piyathasan et al. (2011) empirically. Consequently, I test the first proposition of this conceptual model empirically in the scope of this master thesis. In the following I will shortly introduce the conceptual model of Piyathasan et al. (2011) and give reasons for why I decided to test just one proposition of the model. Below you can see the complete conceptual model by Piyathasan et al. (2011):

² Why I pursued four regression models will become clearer in the course of this thesis. At this point I will just shortly list the four different relationships which will be tested in these four models:

Regression model I: social identity (independent variable) → crowdsourcing engagement (dependent variable)

Regression model II: three sub-factors of social identity (independent variables) → crowdsourcing engagement (dependent variable)

Regression model III: social identity, variables which correlated with crowdsourcing engagement (independent, control variables) → crowdsourcing engagement (dependent variable)

Regression model IV: three sub-factors of social identity, variables which correlated with crowdsourcing engagement (independent, control variables) → crowdsourcing engagement (dependent variable)

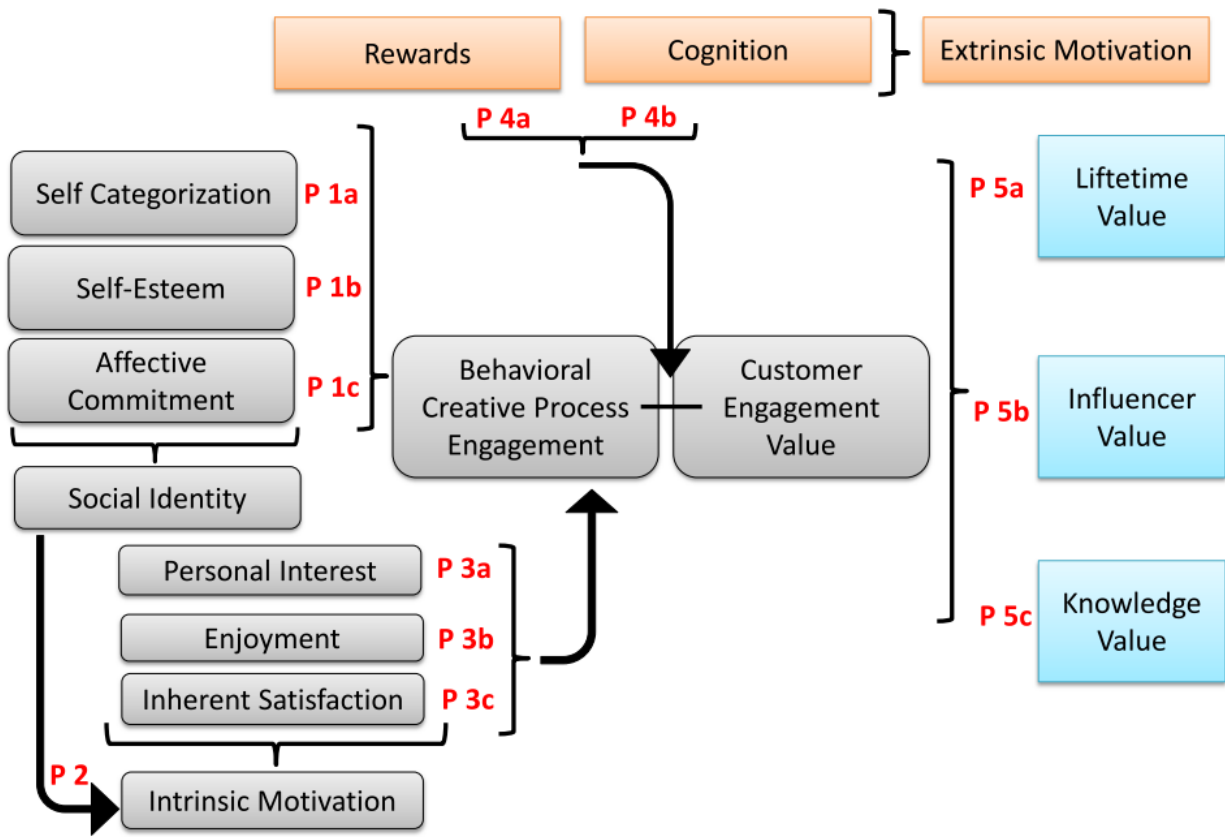


Figure 1: Conceptual model of Piyathasanan et al. (2011) adopted from: (Piyathasanan et al., 2011, p. 1)

The following propositions are displayed in this figure:

P1a – P1c: Greater self categorization, group self-esteem, and affective commitment with the brand increase engagement in the creative crowdsourcing process (Piyathasanan et al., 2011, p. 3).

P2: Social identity with a brand will be positively associated with intrinsic motivation with that brand (Piyathasanan et al., 2011, p. 4).

P3a – P3c: Greater personal interest, enjoyment, and inherent satisfaction to the task increase the engagement in the creative crowdsourcing process (ibid.).

P4a – P4b: Greater attractiveness of rewards and higher expectation of recognition increase value creation from engaging in the creative crowdsourcing process (ibid.).

P5a – P5c: Greater customer engagement in the creative crowdsourcing process increases lifetime value, influencer value, and knowledge value for the firm (Piyathasanan et al., 2011, p. 5).

In the framework of this master thesis I am planning to test the first proposition of the conceptual model empirically and to exclude the remaining four propositions from my analysis. However, it would also make sense to include intrinsic motivation in my empirical analysis and with that also consider the second and the third proposition. Although including intrinsic motivation would be in line with the overarching topic of this thesis which is the question what motivates people to contribute to crowdsourcing I think that handling social identity as one huge theoretical concept of the first proposition in this thesis is sufficient and comprehensive enough. The results of the literature review support this notion when showing that the social identity people perceive is somehow connected to their behavior. However, none of the authors explicitly tested the influence of social identity with its three sub-factors on crowdsourcing engagement empirically. Consequently, I assume that testing the influence of the three factors of social identity on crowdsourcing engagement is complex enough for one thesis and therefore suffice for this graduation project. That is why I neglect the concept of intrinsic motivation and not test the second and third proposition of the model. Additionally, I disregard the last two propositions. Why considering the fourth and fifth proposition would not fit into the concept of this thesis is shortly explained in the following lines: The propositions four and five are concerned with the question what kind of value crowdsourcing initiatives bring about for the respective company. Testing these propositions empirically would include a close cooperation with a company where crowdsourcing projects have been done previously or will be done in the near future. A long-term monitoring program of the crowdsourcing process in at least one company would be needed to draw conclusions about the value crowdsourcing brings about for firms which are actively involved in crowdsourcing. As this short explanation of the last two propositions in the conceptual model has shown these propositions cover a completely different aspect for research than the first part of the model which exclusively deals with the motivation of participants in crowdsourcing projects. Whereas the fourth and the fifth proposition are neglected because they do not match the topic of this thesis, the second and the third proposition is not considered because discussing them would exceed the scope of this thesis. Consequently, only the following proposition is tested empirically in the course of this thesis as a hypothesis:

Greater self categorization, group self-esteem, and affective commitment with the brand increase engagement in the creative crowdsourcing process.

3. Literature review

3.1. Crowdsourcing

In order to fully understand the meaning of the hypothesis a few terms need to be defined and marked off each other. Firstly, to grasp what engaging into a creative crowdsourcing process really means, the term crowdsourcing should be defined. Since Jeff Howe (2006a) originally coined the term “crowdsourcing” the explanation of its meaning starts with Howe’s definition: “Crowdsourcing represents 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 [via the internet – own annotation].” (Howe, 2006a). Whitla (2009) further explains that in order to differentiate the term crowdsourcing from related concepts as ‘wikinomics’ or ‘commons-based peer production’ the condition that companies or organizations offer some kind of reward for outsourcing tasks to the crowd was added later by Howe (Whitla, 2009, pp. 15-16). Additionally, Brabham (2013) adds that the term crowdsourcing is structurally not comparable with peer production because crowdsourcing always has the notion of a top-down format which is directed by the initiating firm or organization (Brabham, 2013). However, Brabham (2013) further mentions that crowdsourcing is also characterized by the sharing of power between the public and the initiating organization (ibid.). He characterizes this as a shared process of “(...) bottom-up open creation by the crowd and top-down management (...)” (Brabham, 2013). Howe’s (2006c) article in the magazine “Wired” indicates that the outsourced tasks or problems by a company do not necessarily need to be solved by the crowd as a whole, but could also be solved by a single person who does not cooperate with the rest of the crowd (Howe, 2006c, pp. 2-3). Howe (2006b) further specifies that crowdsourcing necessarily implies that the company identifies the best suggestion and also later implements this solution (Howe, 2006b). For example, in the case of a design contest this would involve a mass fabrication of that particular design which the company found most convincing (ibid.).

In accordance with Howe’s (2006a) definition crowdsourcing can be classified as a type of open innovation in the sense of Chesbrough (2003). However, according to Guittard and Schenk (2011) a few differences between the concept of open innovation and crowdsourcing should be mentioned here: Firstly, open innovation mostly focuses on innovative processes, whereas crowdsourcing does not necessarily have to

be innovative³ (Schenk & Guittard, 2011, p. 96). Secondly, it is argued that open innovation mainly concentrates on the exchange of knowledge between firms (ibid.). Lichtenthaler (2011) adds to that point that open innovation “(...) considers the trend toward interorganizational innovation processes (...)”⁴ (Lichtenthaler, 2011, p. 77). In another article these external sources of knowledge which firms use to cooperate and exchange with when following the open innovation approach are clearly defined: Enkel et al. (2009) state that companies mostly use clients as their external knowledge source followed by suppliers, competitors and public or commercial research institutes (Enkel et al., 2009, p. 312). Conclusively, this short comparison of different statements within the existing literature shows that it is not clear whether open innovation is restricted only to cooperation between firms or if this concept also includes knowledge exchange with other external organizations or even single persons. Crowdsourcing, in contrast, clearly emphasizes the relationship between a firm and the crowd respectively individuals within this crowd (Schenk & Guittard, 2011, p. 96). Consequently, crowdsourcing can be seen as a way for companies to profit from external knowledge flows (the crowd) without the compelling need to come up with innovative ideas on the basis of these inputs (ibid.).

After having defined the meaning of the term crowdsourcing itself it should be explained to the reader what Piyathananan et al. (2011) understand under a creative crowdsourcing process. According to Piyathananan et al. (2011) the creative crowdsourcing process “refers to the sequence of thoughts and actions of an online community or crowd involved in creativity-relevant tasks, methods or processes and leading to novel and useful ideas to the firm.” (Piyathananan et al., 2011, p. 2). This explanation of a creative crowdsourcing process by Piyathananan et al. (2011) shows that the authors consider crowdsourcing as closely related to Chesbrough’s (2003) definition of open innovation. This is due to the fact that Piyathananan et al. (2011) assume that at the end of the creative crowdsourcing process the members will come up with ideas which are novel and useful for the company which initiated the crowdsourcing tasks. This assumption addresses the possibility of an innovative outcome of a creative

³ It will be shown later that Piyathananan et al. (2011) assume that an innovative idea stands at the end of the crowdsourcing process. The question when an idea emerging from a crowdsourcing process is an innovative one is another point which will not be discussed in the scope of this thesis.

⁴ However, Lichtenthaler (2011) does not clearly define what he means by “interorganizational”. At least it can be concluded that the word “interorganizational” excludes the possibility that open innovation can occur between a company and a sole individual person. Nevertheless, the term “interorganizational” does not clearly imply that these organizational forms necessarily need to be firms as stated by Guittard and Schenk (2011).

crowdsourcing process which parallels the idea behind open innovation. According to Piyathasanan et al. (2011) the creative crowdsourcing process starts when the firm decides that a task which is usually tackled with the help of internal firm resources will be now given to an undefined crowd via the web (ibid.). As a next step several members of the crowd are offered the possibility to perform this specific task in a predefined time frame (ibid.). After each member who tried to solve this task has submitted his or her solution the company will check the quality of work or give it again to the crowd where all members will vote for their favorite solution (ibid.). Finally, the company will use the winning idea which has been identified either by the company itself or – if the company decided to let the crowd vote for the best suggestion – by the crowd for its own benefit (ibid.).

In the following I apply a rather broad conception of the term crowdsourcing and the connected creative crowdsourcing process. I argue that crowdsourcing does not exclusively imply the mere delivery of answers by external actors to the respective companies where problems emerged. It rather also comprises a close cooperation and interaction within the group of external actors (the crowd) as well as with the companies in question. The aspect of close collaboration and interaction within the crowd is in line with the definition of crowdsourcing as stated by Howe (2006a) when he says that crowdsourcing tasks can be solved collaboratively (Howe, 2006a). The fact that crowdsourcing may not only involve interaction among members of the crowd but also include a certain degree of cooperation with the company which initiated the task becomes clear when looking at the creative crowdsourcing process as described by Piyathasanan et al. (2011): After the task has been opened up to the crowd and members of the crowd have submitted possible ideas for the solution to the problem the company may either choose the best suggestion by itself or fall back on people in the crowd by consulting them and asking them to vote for their favorite idea (Piyathasanan et al., 2011, p. 2). Consequently, the top-down structure which was mentioned by Brabham (2013) only concerns the very first step in the crowdsourcing process where a company poses a problem to the crowd. Afterwards, the hierarchical borders between the crowd and the initiating firm get blurred and dissolve into a form of cooperative interactions. This is also in line with what Brabham (2013) later describes in his book as a power-sharing crowdsourcing process (Brabham, 2013). With that he describes the fact that the initiating firm mainly has a management function whereas the crowd pursues an open creation (ibid.). Consequently, in the course of this thesis I apply the following definition of crowdsourcing:

*Crowdsourcing can be defined as a collaborative act of solving tasks within a community in close cooperation with the company which initiated the task in the form of an open call via the internet.*⁵

3.2. Social identity and its relationship to crowdsourcing engagement

In the next lines the term social identity together with its three sub-factors self categorization, group self-esteem and affective commitment with the brand needs to be defined and explained. Additionally, with the help of the literature it will be elucidated to which extent social identity together with its three sub-factors influences people's behavior. In this context it will be shown why it is interesting and necessary to empirically test the relationship between social identity and crowdsourcing engagement.

Defining the term social identity

In their paper Piyathasanan et al. (2011) adopt Tajfel's (1978) definition of the whole concept of social identity (Piyathasanan et al., 2011, p. 2): Social identity is the "part of an individual's self concept which derives from his knowledge of his membership of a social group (or groups) together with the value and emotional significance attached to that membership." (Tajfel, 1978 p. 63). Luhtanen and Crocker (1992) add that social identity further refers to the way people then perceive the social groups they assign themselves to (Luhtanen & Crocker, 1992, p. 302). If the individual now sees himself as a member of a specific social group he will at the same time perceive himself as directly connected and involved in the fate of that particular group and with this he will personally undergo and share successes and failures of this group (Ashforth & Mael, 1989, p. 34; Fred Mael & Tetrick, 1992, pp. 104-105). This fits to one of the main assumptions of social identity theory which is that social identity mainly derives from group membership (Brown, 2000, pp. 746-747). Ellemers et al. (1999) argue that the more people identify with a certain group the more they will behave in terms of their group membership (Ellemers et al., 1999, p. 372). Also Bergami and Bagozzi (2000) agree to that when they explain that the degree of social identity "(...) affects people's tendency to behave in terms of their organization membership (...)" (Bergami & Bagozzi, 2000, p. 574). It is generally assumed that people are striving for a positive social identity (Brown,

⁵ In accordance with this definition and understanding of crowdsourcing it is possible to consider articles about cooperative communities of any kind when examining the influence of social identity on people's behavior with the help of the literature.

2000, p. 747; Trepte, 2006, p. 255). Social identity theory states that people perceive their social identity as positive if they see themselves in connection to the social group they belong to in a good light (Brown, 2000, p. 755). In order to achieve a positive social identity people try to act in favor of the social group they belong to (Trepte, 2006, p. 256). This may be expressed in the way that people show solidarity for the social groups they are attached to (ibid.). Bagozzi and Dholakia (2006) state that people gain a social identity by the perception of three things: Self-awareness of one's membership in a social group, the emotional attachment to this social group and the evaluation of the group membership (Bagozzi & Dholakia, 2006, p. 48). On the basis of this Ellemers et al. (1999) propose that three components may account for people's social identity: "a cognitive component (a cognitive awareness of one's membership in a social group – self categorisation), an evaluative component (a positive or negative value connotation attached to this group membership - group self-esteem), and an emotional component (a sense of emotional involvement with the group - affective commitment)" (Ellemers et al., 1999, p. 372). Piyathasanan et al. (2011) adopt this in their article and also consider the three terms self categorization, group self-esteem and affective commitment as the three components of social identity (Piyathasanan et al., 2011, p. 3). The exact meaning of each of the three sub-factors of social identity will be explained in the next paragraphs. In order to summarize the last paragraph I now give a condensed definition of the term social identity which will be valid throughout this thesis:

Social identity has a cognitive, evaluative and emotional component and therefore contains all the knowledge, perceptions and emotions people receive from their membership in a certain group.

The influence of social identity on people's behavior

In order to follow why social identity might influence people's behavior and with that probably also their crowdsourcing engagement as proposed by Piyathasanan et al. (2011) firstly a paper written by Wang and Tai (2011) was chosen for the following reason: The theoretical concepts the authors present show that they also suppose a connection between the whole concept of social identity and the continual participation of people in communities. Referred to my study such a continual participation could for example take the form of engaging in crowdsourcing projects. Wang and Tai (2011) investigate in their conceptual paper how social presence⁶ influences the continual participation of virtual community

⁶ Definition "social presence": Social presence is "(...) an individual's ability to demonstrate his / her state of being in a virtual environment and so signal his / her availability for interpersonal transactions" (Wang & Tai, 2011, p. 111).

members (Wang & Tai, 2011, p. 110). In order to answer their research question Wang and Tai (2011) propose two conceptual models: The first model is supposed to investigate how social presence influences social identity and with that contributes to the sense of virtual community⁷ which will further lead to satisfaction with the virtual community and a continual participation within the virtual community (Wang & Tai, 2011, p. 112). With their second model Wang and Tai (2011) would like to examine if social presence in virtual communities leads directly to social identity or if the relationship is rather indirectly and goes through the value perception of participating in this community (Wang & Tai, 2011, pp. 112-113). Additionally, the second model also covers the part where the authors ask if an enhanced sense of virtual community leads to a higher satisfaction and with that also to a higher continual participation in the virtual community (ibid.). The difference between both models is the point that in the second model the sense of virtual community cannot only be enhanced by a strong social identity, but also by the perceived value people get from participating in virtual communities (ibid.). Especially the presentation of the first model by Wang and Tai (2011) shows that the authors theoretically propose a relationship between people's social identity they derive from a membership in a certain community and their participation in this community. However, this theoretical departure has not been tested empirically by Wang and Tai (2011) and has also not been applied to crowdsourcing engagement, yet. Nevertheless, their paper shows that the overall theoretical implication of the first proposition by Piyathasanan et al. (2011) is also discussed by other authors.

Secondly, I included a paper written by Blader and Tyler (2009) to show the possible influence of social identity on people's behavior in general. In their paper the authors found empirical significance for the fact that social identity is related to extra-role behavior (Blader & Tyler, 2009, p. 452; 454). Generally, the study deals with the relationship between social identity and extra-role behavior of employees in respect to their work group as well as the whole organization they are employed in (Blader & Tyler, 2009, p. 445). The authors refer to extra-role behavior as the degree to which people carry out work or tasks which go beyond their area of responsibility (ibid.). One can imagine that crowdsourcing could be such a task which is not connected to the usual work requirements of employees. Although my study deals with groups of people who are not necessarily employees of the organization in question it seems to be reasonable to include the study of Blader and Tyler (2009) in this literature review because it is concerned with the way

⁷ Definition "sense of virtual community": Sense of virtual community is "(...) the individual's feelings of membership, influence, and immersion toward a virtual community" (Wang & Tai, 2011, p. 112).

social identity people can influence people's behavior towards this specific organization. In the theoretical section of their article the authors propose that people with a strong social identity towards a certain group are more likely to be interested in the success of this group and would rather act in terms of this group's interest (Blader & Tyler, 2009, p. 446). Consequently, the authors state that the social identity of employees will be positively related to their extra-role behavior (ibid.). In the course of their empirical investigations Blader and Tyler (2009) find that social identity is indeed significantly related to extra-role behavior (Blader & Tyler, 2009, p. 452; 454). As a conclusion one can say that the results of the study by Blader and Tyler (2009) help to support the general underlying concept of this master thesis. This is because the authors have shown that social identity which is derived from group membership positively influences people's behavior in favor of the social group they belong to. From this one could conclude that people who perceive a high social identity with their social group are more likely to participate in a crowdsourcing task than people with a lower degree of social identity. However, the study by Blader and Tyler (2009) does not fully support the first proposition of Piyathasanan et al. (2011). The reasons for that are twofold: Firstly, Blader and Tyler (2009) did not directly focus their study on people's willingness to participate in crowdsourcing projects. Secondly, although the authors used the concept of social identity they only focused on the two sub-factors self categorization and group self-esteem (Blader & Tyler, 2009, p. 448).

As can be concluded from the review of these two articles the idea that people's social identity positively influences their behavior towards the organizations or companies they obtain their social identity from is also supported by other authors. However, in none of the reviewed papers the authors explicitly tested the influence of the three sub-factors of social identity on crowdsourcing engagement. That is why it seems to be reasonable to adopt the following proposition as a hypothesis and to test it empirically in the course of this thesis:

H1: Greater social identity consisting of the sub-factors self categorization, group self-esteem, and affective commitment with the brand increase engagement in the creative crowdsourcing process.

Definition of self categorization

After the general understanding of social identity in the literature has been explained and its possible effect on people's behavior has been introduced it is now of interest to clarify the meanings of its three

sub-factors and also their connection to people's behavior. As a first step self categorization as the first sub-factor of the overall concept of social identity is elucidated. First of all it should be mentioned that in the following paragraph the two terms self categorization and social or organizational identification are used interchangeable. This approach is in accordance with Bergami and Bagozzi (2000) who clearly state in their paper that both terms mean the same (Bergami & Bagozzi, 2000, p. 555). Consequently, I also included papers in this section where the authors used the term social or organizational identification instead of self categorization. Self categorization is the cognitive component and can be described as the process of forming self-categories of organizational membership (Piyathasanan et al., 2011, p. 3). In simple words this process describes how people perceive an aggregation of other people - including themselves - as a group and also realize the connected consequences (see: Haslam, 1997). As stated by several authors people identify with an organization if they perceive some kind of oneness with the organization and further define themselves in terms of their membership in this organization (Bergami & Bagozzi, 2000, p. 557; Fred Mael & Ashforth, 1992, p. 109; Fred Mael & Tetrick, 1992, p. 814). Mael and Ashforth (1992) argue that organizational identification describes the process in which people define their own personality in terms of their membership to a certain group (Fred Mael & Ashforth, 1992, p. 105). Or as Blader and Tyler (2009) explain it the concept self categorization "(...) captures the extent to which group membership is self-defining and thus determines the implications of group membership for how people think about themselves." (Blader & Tyler, 2009, p. 448). Generally, one could say that organizational identification implies that people have the feeling of being intertwined with the social group they belong to and therefore experience a strong emotional tie (Fred Mael & Tetrick, 1992, p. 813). Knippenberg and Schie (2000) go one step further by saying that with organizational identification people not only define themselves as a part of a specific organization but also adopt characteristics that are typical for the organization in question (Knippenberg & Schie, 2000, p. 138). As a consequence people tend to change their behaviors and attitudes in favor of the social groups they belong to (ibid.). In other words this means the following: "The more an individual identifies with an organization, the more likely he or she is to take the organization's perspective and to act in the organization's best interest." (Knippenberg & Schie, 2000, p. 138). Riketta (2005) nicely summarizes the points all the definitions of organizational identification have in common when she writes that people identify themselves with an organization if they link their self-concept to the organization they belong to (Riketta, 2005, p. 361). This usually happens in a cognitive form where individuals identify themselves as a part of the organization and internalize its values (ibid.). Also Dukerich et al. (2002) agree to this definition of organizational identification and add that they consider organizational identification as strong if members embrace the values and goals of the organization in the

way that they adopt these for themselves (Dukerich et al., 2002, pp. 508-509). In this context Knippenberg and Sleenbos (2006) introduce the concept of collective interest which is the incorporation of interests, values and norms of an organization into the self-concept (Knippenberg & Sleenbos, 2006, p. 27). As a consequence, individuals perceive the collective interest as their self-interest and are therefore intrinsically motivated to act on behalf of the collective (ibid.). The following definition of the term self categorization is the attempt to summarize this paragraph as well as to give an own definition of the term on the basis of the literature:

Self categorization is the cognitive component of social identity and describes the awareness of the process of merging one's own personality with a group identity.

The influence of self categorization on people's behavior

To shed some light on the connection between self categorization and people's behavior towards or engagement in the community a paper by Zaglia (2013) was selected which presents the concept of brand communities⁸ (Zaglia, 2013, p. 216). In the beginning of her article Zaglia (2013) firstly highlights the impact brand communities can have: "Social interactions between community members profoundly influence customers' relationship with, and attitude towards, the brand." (ibid.). Additionally, brand communities are also beneficial for the companies in question by helping them to deepen their customer relationships (ibid.). Here especially online computer mediated forms of communication have an accelerating effect on this development (ibid.). Furthermore, Zaglia (2013) mentions three common characteristics of brand communities: Consciousness of kind⁹, shared rituals and tradition and moral responsibility (Zaglia, 2013, pp. 217-218). Consequently, a look at these three characteristics of brand communities helps to understand why people who are deeply rooted in such a community are more likely to support their own community than any other organization. This is the case because these members feel closely connected to their brand community, participate in its tradition and rituals and also feel morally responsible for this community (ibid.). The author found that identification with a brand is one

⁸ Definition "brand community": Brand communities are a special form of consumer communities where consumers interact with other consumers (Zaglia, 2013, p. 216).

⁹ Definition "consciousness of kind": Consciousness of kind describes the perceived membership of participants in brand communities (Zaglia, 2013, p. 217). Usually, members feel connected to other members of the community and separate from outsiders (in-group and out-group comparison) (ibid.). This term is related to social identity theory (ibid.).

reason for an enforced motivation to actively participate in this brand community (Zaglia, 2013, p. 221). Participation could for example be engaging in discussions via the online platform of this brand community (ibid.). This finding is closely related to what Piyathananan et al. (2011) state in their first sub-proposition when they say that people's identification with an organization or group leads to an increased crowdsourcing engagement. Nevertheless, the study by Zaglia (2013) does not entirely support the theoretical assumption in the first sub-proposition of Piyathananan et al. (2011) because it exclusively considers brand communities and does not draw any conclusions to other communities such as crowdsourcing communities which are relevant for my study.

A second study which clarifies the relationship between self categorization and people's behavior was written by Woisetschläger et al. (2008). It empirically proves that the identification with an organization influences people's behavior towards this particular organization (Woisetschläger et al., 2008, p. 252). Woisetschläger et al. (2008) examined the drivers of consumer brand community participation. The authors assume that a company can only establish a successful brand community of itself if the community members adequately participate in it (Woisetschläger et al., 2008, pp. 241-242). In order to tackle their research issues the authors used social identity theory as a basis for their theoretical framework (Woisetschläger et al., 2008, p. 240). In doing this they propose that the identification with the brand community will have a positive influence on the willingness of people to actively participate in these communities (Woisetschläger et al., 2008, p. 243). Finally, the results of their empirical analysis show that identification with the community is one of the main drivers which explains most of the variance in community participation (Woisetschläger et al., 2008, p. 252). Consequently, this study empirically proved one key proposal of social identity theory which is that how people identify with an organization influences their behavior towards this organization (Ellemers et al., 1999, p. 372). Conclusively, this study at least partly supports the theoretical idea behind the first sub-proposition of Piyathananan et al. (2011). I chose the expression "partly supports" on purpose because the study at hand did not directly deal with the influence of group identification on people's crowdsourcing engagement but rather remained more general when only considering the overall participation of people in brand communities.

The review of these two articles displays that an increased self categorization has the potential to increase people's participation in any form. Anyhow, the influence of an increased self categorization on crowdsourcing engagement has not been empirically tested, yet. That is why it is still worthwhile to adopt

the following first sub-proposition by Piyathananan et al. (2011) as a first sub-hypothesis and to test it empirically in this thesis:

H1a: Greater self categorization with the brand increase engagement in the creative crowdsourcing process.

Definition of self-esteem

Before the second sub-factor of the construct social identity will be further discussed it should first be mentioned that Piyathananan et al. (2011) use group self-esteem as well as organization-based self-esteem to name the second factor of social identity (Piyathananan et al., 2011, p. 3). Furthermore, I also included scientific articles which used the expression collective self-esteem to describe the second sub-factor of social identity. It is assumed that all three nomenclatures appeal to the fact that the type of self-esteem which is relevant for the concept of social identity derives from the membership in a social group. That is why I consider all three terms as interchangeable. Group self-esteem is the second component of social identity and is called the evaluative component (Piyathananan et al., 2011, p. 3). Bergami and Bagozzi (2000) define organization-based self-esteem as follows: Organizational-based self-esteem refers to the "(...) evaluations of self-worth deriving from one's membership in the organization." (Bergami & Bagozzi, 2000, p. 560). During this evaluation process people also consider recent successes or failures of the group they are part of (Gangadharbatla, 2008, p. 8). For example, if the group just experienced a considerable loss in any form the self-worth of members of this group would be lowered (ibid.). Depending on the fact whether this evaluation derived from group membership is positive or negative also the perceived social identity will be positive or negative (Luhtanen & Crocker, 1992, p. 304). Consequently, the value people attach to their group membership and its influence on people's social identity in turn also impacts the way people feel and think about themselves (Blader & Tyler, 2009, p. 448). If people attach positive feelings to a group and subsequently are able to identify themselves with this group they will try to act in the interest of the group (Piyathananan et al., 2011, p. 3). According to Gangadharbatla (2008) people generally strive for a high level of collective self-esteem (Gangadharbatla, 2008, p. 8). The next sentence summarizes the concept of self-esteem¹⁰ in the form of an own definition which is supposed to serve as a basis in the course of this thesis:

¹⁰ In the following the terms self-esteem, group self-esteem, organization-based self-esteem and collective self-esteem will be used interchangeable and all refer to the second sub-factor of social identity.

***Self-esteem** is the evaluative component of social identity and entails an evaluation process of the own personality against the background of the group membership.*

The influence of self-esteem on people's behavior

For understanding the relationship between organization-based self-esteem and people's behavior towards the organization they acquire their self-esteem from an article written by Pierce et al. (1989) is introduced: In their paper the authors explore the whole theoretical concept of organization-based self-esteem by defining the term and also giving recommendations how to measure it (Pierce et al., 1989, p. 622). At the beginning of their article the authors state the proposition that people with a high self-esteem are also more likely to perform at a high level in the organization or group they belong to (Pierce et al., 1989, p. 623). As Pierce et al. (1989) later explain people with a high organization-based self-esteem think about themselves as a meaningful part of the organization (Pierce et al., 1989, p. 625). Consequently, they are also more persuaded that they are able to handle tasks they are supposed to fulfil than people with a lower self-esteem (ibid.). In the course of their paper the authors additionally develop the thought that people with high organization-based self-esteem are likely to engage in favor of the organization they belong to (Pierce et al., 1989, p. 630). Pierce et al. (1989) explain this not only by the fact that the individuals featuring a high self-esteem are confident to conduct various tasks for the organization they belong to but also by the matter that these people are eager to support their organization (ibid.). During their empirical investigations the authors found support for all of their theoretical constructs they developed in their paper (Pierce et al., 1989, p. 644). Summarizing, one can say that the study by Pierce et al. (1989) has shown that people's self-esteem influences their willingness to behave favorable for the organization in question. For my study this could mean that people with high self-esteem are more likely to engage in crowdsourcing for the organization they belong to and with that doing their organization a favor in that they participate in crowdsourcing tasks for their organization. Consequently, the study by Pierce et al. (1989) serves as an interesting starting point for further investigations. However, it is not clear if the findings about the relationship between self-esteem and people's behavior towards the organization can also be applied to a crowdsourcing environment.

Furthermore, a paper in the form of an extensive literature review which deals among other things with the effects of organization-based self-esteem by Pierce and Gardner (2004) is helpful. Pierce and Gardner

(2004) start their paper with the fact that employee self-esteem which is derived from daily experiences at work and its consequences on employee motivation, work-related attitudes and behaviors started to become a research issue back in the 1970s (Pierce & Gardner, 2004, p. 591). The authors cite nine studies published between 1989 and 2003 which empirically confirmed that organization-based self-esteem positively influences people's performance at their working place (Pierce & Gardner, 2004, pp. 608-609). Consequently, the authors draw the conclusion that the literature has clearly shown that organization-based self-esteem is positively related to productive working behaviors such as a high performance (Pierce & Gardner, 2004, pp. 609-610). The question which remains is whether it is possible to transfer the findings reported by Pierce and Gardner (2004) from the work environment into the crowdsourcing context.

After the discussion of these two scientific papers it becomes clear that the relationship between self-esteem and crowdsourcing engagement needs to be tested empirically. This is important in order to know whether the connection between self-esteem and people's behavior towards the organizations they belong to found in the literature also holds true in a crowdsourcing context. Consequently, it is reasonable to adopt the following second sub-proposition by Piyathasanan et al. (2011) as a second sub-hypothesis and to test it empirically in the course of this thesis:

H1b: Group self-esteem increase engagement in the creative crowdsourcing process.

Definition of affective commitment

The third factor of social identity, affective commitment, is the emotional component of social identity (Piyathasanan et al., 2011, p. 3). According to Meyer et al. (2002) affective commitment implies an emotional attachment to the organization and an involvement in the organization (Meyer et al., 2002, p. 21). Allen and Meyer (1990) say that an individual is affectively committed to a certain group if he or she feels a strong attachment to that group and enjoys his or her membership in that particular group (Allen & Meyer, 1990, p. 2). Buchanan (1974) adds to that that if people are strongly committed to a certain organization they will also take up the values and goals of that organization and internalize them (Buchanan, 1974, p. 533). This fits to the fact that Fullerton (2003) reports that affective commitment represents the desire to foster a precious relationship (Fullerton, 2003, p. 334). Consequently, the sources of affective commitment are the following: Identification, shared values, belongingness, dedication and

similarity (ibid.). Bergami and Bagozzi (2000) explain further that if people feel attached to a group they derive joy and love from their group membership (Bergami & Bagozzi, 2000, p. 560). In other words if an individual is strongly involved in an organization and also enjoys this membership in the organization one can speak of affective commitment with an organization (Allen & Meyer, 1990, p. 2). Or as Fullerton (2003) puts it affective commitment “(...) represents an enduring desire to maintain a valued relationship.” (Fullerton, 2003, p. 334). Culpepper (2000) further mentions three factors of affective commitment: The acceptance of the organization’s goals, the willingness to come up with an immense effort for the organization and the wish to remain a member of this organization (Culpepper, 2000, p. 605). Ellemers et al. (1999) clearly point out that the degree to which people perceive an affective commitment to a certain group positively influences the way people act in terms of their membership in this group (Ellemers et al., 1999, p. 385). This is also in accordance with the argumentation of Wetzels et al. (1998) when they say that a strong affective commitment positively influences people’s intention and desire to remain in the relationship with the respective group and their performance in this group as well as their willingness to invest in this relationship (Wetzels et al., 1998, p. 409). Gardner et al. (2011) add to that that people with a strong affective commitment will behave in a way which clearly benefits the organization they derive their affective commitment from (Gardner et al., 2011, p. 317). In accordance with that Piyathananan et al. (2011) assume that if affective commitment of a person to the brand is high this person will also show character traits such as altruism, sportsmanship and civic virtue (Piyathananan et al., 2011, p. 3). In the next sentence the essence of the term affective commitment is summarized with the help of an own short definition:

Affective commitment is the emotional component of social identity and reflects whether people enjoy their attachment to the group and if they are willing to act in favor of their group.

The influence of affective commitment on people’s behavior

For understanding the relationship between affective commitment and people’s willingness to engage in activities for the organization they are a member of a paper written by Cheung and Lee (2012) provides helpful insights. In their study the authors confirmed empirically that an emotional connection to a group increases people’s willingness to engage for this particular group (Cheung & Lee, 2012, p. 222). In the case of my study such an increased engagement could be expressed by an increased willingness to take part in

crowdsourcing projects. In their paper Cheung and Lee (2012) found that sense of belonging¹¹ had the highest impact on consumer's intention to actively take part in online opinion platforms (Cheung & Lee, 2012, p. 222). The authors explain these findings with the fact that people who feel closely connected and attached to a certain group are likely to engage for the well-being of this community in any form (ibid.). Therefore, sense of belonging is a form of collective motivation where the benefit of the whole group is more important than the individual return (Cheung & Lee, 2012, p. 220). The concept of collective motivation is in turn closely related to social identity theory which states that individuals gain identity from the group they belong to (ibid.). Consequently, if people feel strongly connected to a certain group (high sense of belonging) they are likely to define themselves in terms of their membership to this particular group (social identity) and are with that also more likely to engage for the sake of this group (collective motivation) (ibid.). Due to the fact that the authors explicitly state that their definition of sense of belonging can be seen as totally equivalent to the term affective commitment (ibid.) it can be argued that this paper at least partly supports the theoretical idea behind the third sub-proposition by Piyathasanan et al. (2011). However, the study by Cheung and Lee (2012) does not fully support what Piyathasanan et al. (2011) state in their third sub-proposition because it does not consider people's engagement in crowdsourcing but rather concentrates on participation in online consumer-opinion platforms.

Additionally, also a study written by Den Hartog and Belschak (2007) is interesting because the authors empirically prove that affective commitment influences employee's tendencies to come up with initiatives at their work places (Den Hartog & Belschak, 2007, p. 610). The authors researched – among other issues – the relationship between employees' affective commitment at work and their connected initiatives they come up with during daily work (Den Hartog & Belschak, 2007, p. 601). Den Hartog and Belschak (2007) define initiative as proactive activities which support the goals of the respective organization (Den Hartog & Belschak, 2007, p. 605). For instance, this may take place in the form of making suggestions for improvements, helping other employees at work or solving internal problems which might not be explicitly assigned to the person working on it when taking up an initiative (ibid.). Engagement in a crowdsourcing task could be counted as taking an initiative on behalf of a certain organization. After having interviewed several employees of large hospitals in the Netherlands the authors empirically prove

¹¹ Definition "sense of belonging": "Sense of belonging refers to a sense of emotional involvement with the group." (Cheung & Lee, 2012, p. 220).

that affective commitment positively influences employees' tendencies to come up with initiatives (Den Hartog & Belschak, 2007, p. 607; 610). This study fits nicely into the theoretical context of this thesis because the results support to a certain degree the third sub-proposition by Piyathasanan et al. (2011) which is that affective commitment positively affects people's tendency to engage in crowdsourcing initiatives of the group they derive their social identity from. Here, one could argue that engaging in crowdsourcing initiatives somehow resembles the act of coming up with initiatives which was examined in the study. Nonetheless, the study by Den Hartog and Belschak (2007) did only cover the behavior of employees of hospitals and did not directly explore the effect of a high affective commitment on crowdsourcing engagement.

An insight into both papers has shown that there persists a connection between people's affective commitment to a group and their engagement in this particular group. However, the direct influence of affective commitment on crowdsourcing engagement has not been empirically tested so far. Consequently, it is plausible to take the following third sub-proposition by Piyathasanan et al. (2011) as a sub-hypothesis and to test it empirically:

H1c: Greater affective commitment with the brand increase engagement in the creative crowdsourcing process.

This literature review shows that all the concepts covered in the first proposition and its three sub-propositions by Piyathasanan et al. (2011) are of interest in many publications until today. However, it also revealed that none of the studies explicitly tested the influence of social identity with its three sub-factors on people's crowdsourcing engagement. Consequently, the aim of this study is to fill this gap in the literature. Therefore, in this study the following hypothesis with its three sub-hypotheses is tested empirically:

H1: Greater self categorization, group self-esteem, and affective commitment with the brand increase engagement in the creative crowdsourcing process.

H1a: Greater self categorization increase engagement in the creative crowdsourcing process.

H1b: Greater group self-esteem increase engagement in the creative crowdsourcing process.

H1c: Greater affective commitment with the brand increase engagement in the creative crowdsourcing process.

4. Method

After having highlighted the conceptual background and the objection of this thesis it is now time to introduce the methodology which is used to answer the research hypothesis and its sub-hypotheses. In this section firstly the research setting is described. Here, the reader will get to know why the format of an online questionnaire was chosen and where the questionnaire was spread. Secondly, the sample at hand is described and its characteristics are presented. Thirdly, it is explained how the variables used in this study were operationalized. For this firstly, the development and measurement of all the variables used in this study are exhibited. This is followed by the introduction and description of the control variables for this study. The method part is finished by enumerating the steps which are necessary to analyze the data at hand.

4.1. Research setting

The general idea for testing the first proposition together with its three sub-propositions by Piyathasanan et al. (2011) empirically is to use an online questionnaire which contains questions about the perceived social identity of people to BMW as well as questions about the willingness of survey participants to engage in crowdsourcing activities of BMW. The advantages of such an online questionnaire especially for the scope of a graduating project are twofold: Firstly, it is easy to distribute the link for the survey via mail or social network forums all over the world (Evans & Mathur, 2005, pp. 196-198). Secondly, the possibility of electronically transmitting the data from the online questionnaire platform into a statistic program such as SPSS¹² afterwards will facilitate things not only in terms of ease of use but also in terms of costs (Evans & Mathur, 2005, pp. 198-199). In the appendix the reader will find screenshots of the online questionnaire.

¹² Here, it should be mentioned that I will use SPSS as a statistical software to evaluate my data.

Altogether I received 278 valid responses. Figure 2 shortly summarizes the sample characteristics by listing the average, standard deviation, minimum and maximum values of personal information the study participants were requested about:

4.2. Description of the sample

Characteristic	Average	St. Dev.	Min.	Max.	N
<i>Country of living</i>	1.5216	.50043	1	2	278
<i>Gender</i>	1.4173	.49400	1	2	278
<i>Born in Estonia</i>	.4353	.49668	0	1	278
<i>Born in Switzerland</i>	.4568	.49903	0	1	278
<i>Born in Germany</i>	.0396	.19529	0	1	278
<i>Born in Spain</i>	.0036	.05998	0	1	278
<i>Born in Lithuania</i>	.0180	.13314	0	1	278
<i>Born in Austria</i>	.0288	.16748	0	1	278
<i>Born in Finland</i>	.0180	.13314	0	1	278
<i>Age</i>	1970.219	14.1257	1940	1995	278
<i>High school</i>	.0971	.29666	0	1	278
<i>Abitur</i>	.2986	.45845	0	1	278
<i>Diploma</i>	.2194	.41460	0	1	278
<i>Bachelor</i>	.1475	.35523	0	1	278
<i>Master</i>	.0180	.13314	0	1	278
<i>PhD</i>	.0612	.24004	0	1	278
<i>None of these</i>	.1583	.36566	0	1	278
<i>Student</i>	.1187	.32402	0	1	278
<i>Retired</i>	.0827	.27598	0	1	278
<i>Half-time employed</i>	.2014	.40180	0	1	278
<i>Full-time (self-)employed</i>	.5360	.49960	0	1	278
<i>Unemployed</i>	.0612	.24004	0	1	278

Figure 2: Sample characteristics - summarizing table

Looking at the whole sample one can say that roughly one half (52.2 %, 145 persons) of the study participants are living in Switzerland, whereas approximately the other half (47.8 %, 133 persons) stays in Estonia at the moment. When it comes to the gender of study participants it can be said that most of the persons questioned are men (58.3 %, 162 persons). Almost half of all study participants were born in Switzerland (45.7 %, 127 persons). The second largest group regarding their country of birth are study participants who were born in Estonia (43.5 %, 121 persons). Concerning the age level the sample is quite diverse. Nevertheless, at least it can be said that the largest age group contains people who were born in the year 1975 (4.7 %, 13 persons). This group is closely followed by study participants who were born in

1979 (4.3 %, 12 persons). Moreover, roughly one third of study participants declared the Abitur as their highest level of education (29.9 %, 83 persons). The second largest group is formed by people whose highest educational degree is the Diploma (21.9 %, 61 persons). When it comes to their current status of occupation more than the half of surveyed people are full-time (self-)employed (53.6 %, 149 persons). The two largest groups among the other half of study participants concerning the occupation status consist of half-time employed persons (20.1 %, 56 persons) and students (11.9 %, 33 persons).

4.3. Operationalization of variables

4.3.1. Processing of variables

In order to being able to include all the relevant variables in my empirical analyses it is important to execute several steps first: Firstly, it is necessary to change all of the string variables which only contain text into numeric values in order to be able to proceed with these variables in the course of my statistical analysis. In the case of my study the string variables which I need to recode are gender, country of birth, home country, educational degree, occupation status and previous crowdsourcing participation. Secondly, I have to recode all items which are reversely poled at the moment to ensure that all items are poled in the same way – in this case positively. This step is necessary to simplify the application of other methods later in the course of the evaluation process. Thirdly, I need to recode all the variables which have more than two categories and are not measured at an interval level into dummy variables. This is a necessary step because with the help of dummy coding one of the major assumptions of the regression model is fulfilled. This assumption says that all independent and control variables must be measured at the interval level or categorical with two categories (Field, 2005, p. 169). Due to the fact that when pursuing dummy coding only the numbers 0 and 1 are used the new dummy variables only have two categories (Field, 2005, p. 208). The variables which need to be recoded in dummy variables are the country people are born in, the level of education people possess and people's occupation status. All other variables I use either have only two categories (gender → male or female, country of living → Estonia or Switzerland, previous crowdsourcing participation → yes or no) or are measured at an interval level¹³

¹³ Here, it should be mentioned that it is generally assumed that Likert scales are measured on an interval level (Field, 2005, p. 641).

(age, social identity, perceived capabilities, hope to earn money, firm recognition, hope to enhance professional reputation).

4.3.2. Measuring social identity, crowdsourcing engagement and control variables

In order to measure the three social identity sub-factors I decided to scan the literature for pre-existing scales which already measured each of the factors and attained a high reliability and validity for that. This approach is consistent with what Hyman et al. (2006) describe when they are talking about the use of pre-existing survey questions. They argue that using already developed questions brings several advantages about: Firstly, such questions have already been tested of their usefulness (Hyman et al., 2006, p. 3). Secondly, it can save the researcher time and sometimes money for developing questions (ibid.). Additionally, the authors state that conceptual, methodological and maybe even measurement work which has been done after the publication of the existing questions helps to complement and update them (Hyman et al., 2006, p. 4). Due to the fact that not every question or statement I found in the existing literature was particularly suitable for my study I needed to carry out some adjustments. According to Hyman et al. (2006) this procedure is reasonable (Hyman et al., 2006, pp. 4-5). However, the authors advise other researchers that if they use pre-existing items or even whole scales or questionnaires they need to test them for reliability and validity (Hyman et al., 2006, pp. 6-7). In the case of my study design I took out some items of existing scales because they did not fit into my study for various reasons. The explicit reasons will be mentioned in detail later on. At this point it should just be mentioned that this approach does not seem to be unusual. Bergami and Bagozzi (2000) also adopted only six items of an already existing scale for organizational-based self-esteem which originally contains twenty items (Bergami & Bagozzi, 2000, p. 565). Surprisingly, these authors do not even give reasons for that but only enumerate the six items they picked from the original scale. Considering this approach it could be concluded that excluding items from existing scales may be reasonable. Additionally, in contrast to Bergami and Bagozzi (2000) I will give reasons for why I think that certain items do not particularly fit in my study. In the course of this study the applied scales will be tested for reliability, however, also testing their validity would exceed the scope of this master thesis. Assessing the construct validity of a test would involve proving convergent and discriminant validity (Campbell & Fiske, 1959, p. 81). In order to prove convergent validity one would have to show that applying different methods for evaluating the data leads to the same measurements of the explored construct (ibid.). Testing for discriminant validity would imply to prove that the measurements of the examined construct differ with the measurements of other

constructs (ibid.). As one can conclude from this the process of assessing the validity of the scales used in this study would be too extensive for the scope of a master thesis.

Scale for self-categorization

Original scale for self categorization by Mael and Tetrick (1992)¹⁴	Scale for self categorization used in this study adapted from Mael and Tetrick (1992)
When someone criticizes (this organization), it feels like a personal insult	When someone criticizes BMW, it feels like a personal insult.
I'm very interested in what others think about (this organization).	I'm very interested in what others think about BMW.
When I talk about this organization, I usually say "we" rather than "they".	When I talk about BMW, I usually say 'we' rather than 'they'.
This organization's successes are my successes.	BMW's successes are my successes.
When someone praises this organization, it feels like a personal compliment.	When someone praises BMW, it feels like a personal compliment.
I act like (name of the organization) person to a great extent.	--- not used ---
If a story in the media criticized this organization, I would feel embarrassed.	If a story in the media criticized BMW, I would feel embarrassed.
I don't act like a typical (name of the organization) person. (R)	--- not used ---
I have a number of qualities typical of (name of the organization) people.	--- not used ---
The limitations associated with (name of the organization) people apply to me also.	--- not used ---

Figure 3: Scale for self categorization - partly adapted from (Fred Mael & Tetrick, 1992, p. 817)

The authors Mael and Tetrick (1992) developed a ten-item scale for self categorization (Fred Mael & Tetrick, 1992, p. 817). Six out of ten items were used for this study. This scale of ten items evolved after an extensive literature review about organizational identification and a pilot study with undergraduates regarding their identification with their university (Fred Mael & Tetrick, 1992, pp. 815-816). Additionally, Knippenberg and Schie (2000) confirmed the high reliability of the scale of Mael and Tetrick (1992) (Knippenberg & Schie, 2000, p. 141). Riketta (2005) also proved that the scale constructed by Mael and

¹⁴ Despite the title of the table says "identification with a psychological group (IDPG)" the authors intended to measure organizational identification which is a subset of IDPG (Fred Mael & Tetrick, 1992, pp. 814-815). The term organizational identification in turn can be used interchangeable with "organizational self categorization" (Bergami & Bagozzi, 2000, p. 555). Conclusively, it is reasonable to use this scale by Mael and Tetrick (1992) for an empirical study of the factor of organizational self categorization.

Tetrick (1992) is the most representative measurement of organizational identification (self categorization) concerning its empirical outcomes compared to the scales from other studies dealing with organization identification (Riketta, 2005, p. 368). These findings by Riketta (2005) show that the scale by Mael and Tetrick (1992) is valid¹⁵ (Frankfort-Nachmias & Leon-Guerrero, 2010, p. 16). Furthermore, Riketta (2005) enumerates additional advantages of using the scale developed by Mael and Tetrick (1992): It is a relatively short scale which is widely used and easy to administer (Riketta, 2005, p. 374). Last but not least, Riketta (2005) adds that besides his empirical study also several other research results have demonstrated the reliability and validity of this scale (ibid.). However, due to the fact that I changed the scale by Mael and Tetrick (1992) by shortening it I will need to test for its reliability again later in the course of this study. The remaining items of the ten-item scale by Mael and Tetrick (1992) were left out for the reason that they refer more to activities and qualities needed to actually perform a job in this certain organization. This does probably not apply for most survey participants. It should be mentioned that I changed the expression “organization” to “BMW” in all of the six relevant items. The answer options for these items are given in the form of a five-point Likert scale and range from “strongly disagree” to “strongly agree”. For transforming all of the items in this scale into one variable value for each study participant I added all of the single values of each item and divided this value by the number of items. The reliability analysis for the self categorization scale revealed a Cronbach’s alpha of .760 which shows that this scale possesses a high reliability¹⁶.

Scale for self-esteem

Original scale for group self-esteem by Ellemers et al. (1999)	Scale for group self-esteem used in this study adapted from Ellemers et al. (1999)
I think my group has little to be proud of.	I think BMW has little to be proud of.
I feel good about my group.	I feel good about BMW.
I have little respect for my group.	I have little respect for BMW.
I would rather not tell that I belong to this group.	I would rather not tell that I belong to BMW.

Figure 4: Scale for group self-esteem - partly adapted from (Ellemers et al., 1999, p. 379)

¹⁵ Definition of validity: “Validity refers to the extent to which measures indicate what they are intended to measure.” (Frankfort-Nachmias & Leon-Guerrero, 2010, p. 16). This definition of validity matches the understanding of the so-called “construct validity” which is one sub-type of validity (Moosbrugger & Kelava, 2011, p. 153). In the course of this master thesis this definition of validity will be taken as a basis.

¹⁶ This can be said because according to Field (2005) each value of Cronbach’s alpha which lies above .7 indicates a reliable scale (Field, 2005, p. 668).

I argue that the self-esteem scale by Ellemers et al. (2011) best represents that how people think about a group or organization as a whole influences whether they think about their membership in a positive or negative way. This is perfectly in line with the way how Piyathasanan et al. (2011) understand the concept of self-esteem (Piyathasanan et al., 2011, p. 3). Ellemers et al. (1999) empirically tested their scale and found out that the division of social identity into the three factors with the respective items is acceptable (Ellemers et al., 1999, p. 379). This exactly resembles what one understands under a proper validity of a scale. The expression “my group” was changed to “BMW” to keep up the consistency of the way all the items in the three groups are formulated. It should further be noted that the answer options for the four items in this scale are given in the form of a five-point Likert scale and range from “strongly disagree” to “strongly agree”. In order to summarize all the items of the scale for each study participant into one value I added the values of each item and divided this sum value by the number of items. The reliability analysis revealed a value of .720 for Cronbach’s alpha. This indicates a high reliability of the scale.

Scale for affective commitment

Original scale for affective commitment by Allen and Meyer (1990) ¹⁷	Scale for affective commitment used in this study adapted from Allen and Meyer (1990)
I would be very happy to spend the rest of my career with this organization.	I would be very happy to spend the rest of my career with BMW.
I enjoy discussing my organization with people outside it.	I enjoy discussing BMW with people outside it.
I really feel as if this organization’s problems are my own.	I really feel as if BMW’s problems are my own.
I think I could easily become as attached to another organization as I am to this one. (R)	--- not used ---
I do not feel like ‘part of the family’ at my organization. (R)	--- not used ---
I do not feel ‘emotionally attached’ to this organization. (R)	I do not feel ‘emotionally attached’ to BMW.
This organization has a great deal of personal meaning for me.	BMW has a great deal of personal meaning for me.
I do not feel a strong sense of belonging to my organization. (R)	I do not feel a strong sense of belonging to BMW.

Figure 5: Scale for affective commitment - partly adapted from (Allen & Meyer, 1990, p. 6)

¹⁷ Due to the fact that Piyathasanan et al. (2011) only mention affective commitment as one of the three factors of social identity in their first proposition the other two commitment scales (continuance and normative commitment) by Allen and Meyer (1990) will be neglected.

I think that the scale by Allen and Meyer (1990) is trustworthy and with that usable for this study because the authors proved with an empirical study that affective commitment to an organization can be measured using the items stated above (Allen & Meyer, 1990, p. 13). Additionally, Culpepper (2000) reports that using this scale produces reliable and valid results (Culpepper, 2000, p. 605). Some years later Jaros (2007) states that the scale by Allen and Meyer (1991) is still dominating literature which is concerned with organizational commitment (Jaros, 2007, p. 7). Also Riketta (2005) noticed that the affective commitment scale by Allen and Meyer (1990) is one of the two most used scales to measure affective organizational commitment (Riketta, 2005, p. 361). Conclusively, it can be stated that it is reasonable to use this affective commitment scale for my study. Additionally, it can be argued that the scale by Allen and Meyer (1990) is more appropriate than other possible commitment scales which do not differentiate between different components of commitment because it considers affective commitment as a separate factor besides continuance and normative commitment (Allen & Meyer, 1990, p. 3). This is totally in line with the first proposition of Piyathasanan et al. (2011) where only the affective commitment of individuals to a group is considered as one factor of social identity (Piyathasanan et al., 2011, p. 3). However, I deleted the items four and five due to the fact that this study also addresses people who are not members of BMW. Additionally, I changed expressions as “this organization” or “my organization” to “BMW”. The answer options for the items are given in the form of a five-point Likert scale and range from “strongly disagree” to “strongly agree”. For calculating one value for the whole affective commitment scale per study participant I add all of the single item values and divide this sum by the number of items the scale possesses. The scale for affective commitment features a Cronbach’s alpha of .828 which indicates a high reliability.

Now that it has been shown that all the three sub-scales of social identity feature a high reliability it is tested whether the reliability remains high if all the factors of the different sub-scales are taken together into one overarching social identity scale. With a Cronbach’s alpha of .9 also the huge social identity scale shows a high reliability. Consequently, it is reasonable to work with the three sub-scales as well as with the summarized social identity scale throughout the course of this study. In order to be able to also work with a social identity scale I now created one single social identity value for each of the study participants. This was done by calculating the mean value of all the different variables from the three sub-scales of social identity. The table below summarizes the findings from the reliability analyses of the three sub-scales and the summarized social identity scale:

	Statistics of Reliability for Self Categorization Scale	Statistics of Reliability for Self-Esteem Scale	Statistics of Reliability for Affective Commitment Scale	Statistics of Reliability for Social Identity Scale
Cronbach's Alpha	,760	,720	,828	,900
Number of Items	6	4	6	16

Figure 6: Reliabilities of the scales – self categorization, self-esteem, affective commitment, social identity

Item for creative crowdsourcing engagement

Item for creative crowdsourcing engagement
I would like to engage in creative crowdsourcing projects of BMW similar to this in the future.

Figure 7: Item for creative crowdsourcing engagement

This item was created by myself and is supposed to address the participants' willingness to engage in creative crowdsourcing projects of BMW. In the case of this item the study participants are requested to read some instructions before they are supposed to choose an answer. For the instructions I firstly shortly explained what a crowdsourcing task actually is in my own words. The wording of the first instruction is as follows: "A crowdsourcing task is a task which a firm does not want to handle internally but rather opens up for the public to solve it. By that the firm hopes to find any person or group of people who has a clever solution to the outsourced task which can be implemented later." Secondly, in order to give the participants an impression of an example for an actual creative crowdsourcing task the trunk idea contest by BMW was depicted with the help of some screenshots from the BMW Group co-creation lab. This item for crowdsourcing engagement is necessary for the questionnaire of this study because the aim of this study is to find out whether there is a connection between social identity people derive from a certain organization and people's engagement in creative crowdsourcing tasks assigned by the same organization they have derived their social identity from. Lastly, according to Piyathasan et al. (2011) it should be mentioned that people are engaged in creative crowdsourcing processes as soon as they show interest in the crowdsourcing task in question (Piyathasan et al., 2011, p. 2). Due to the fact that this item directly addresses the interest in a certain type of creative crowdsourcing task it can be said that choosing only this item to measure creative crowdsourcing engagement is sufficient for this study. Furthermore, it can be argued that creative crowdsourcing engagement is simply the willingness to participate in a creative crowdsourcing project. Therefore, creative crowdsourcing engagement can be measured by giving study participants the option to choose to what degree they would be willing to participate in creative crowdsourcing tasks. Consequently, only one item is sufficient to ask study participants if they would like to engage in creative crowdsourcing projects. The answer options for this item are given in the form of a five-point Likert scale and range from "strongly disagree" to "strongly agree".

Control variables

After the sub-scales of social identity and also the item for crowdsourcing engagement have been introduced it is now time to come to the topic of control variables. In their “Dictionary of Statistics and Methodology” for social sciences Voigt and Johnson (2011) define the term “control variable” as follows: A control variable is “an extraneous variable that you do not wish to examine in your study; hence you control for it [...]” (Vogt & Johnson, 2011, p. 76). The authors further explain that this process of controlling can statistically be done with methods as partial correlation or regression analysis (ibid.). In accordance with Frankfort-Nachmias and Leon-Guerrero (2010) this process of controlling for additional variables (control variables) in order to further explore the bivariate relationship between dependent and independent variable is called “elaboration” (Frankfort-Nachmias & Leon-Guerrero, 2010, p. 310). Every possible control variable has the potential to give an alternative explanation of the relationship between both considered variables (ibid.). According to the authors the use of control variables in a bivariate analysis pursues three major goals: Firstly, the process of elaboration helps the researcher to give theoretical and empirical evidence that the relationship between the independent and dependent variable cannot be explained by another causally prior control variable (Frankfort-Nachmias & Leon-Guerrero, 2010, p. 310; 324). In other words the process of elaboration “[...] allows us to test for nonspuriousness” (Frankfort-Nachmias & Leon-Guerrero, 2010, p. 310). Secondly, elaboration clears up the bivariate relationship because new variables are brought up and it is hypothesized that they might intervene in the relationship between independent and dependent variable (ibid.). Thirdly, the introduction of additional control variables “[...] specifies the different conditions under which the original bivariate relationship might be hold.” (ibid.).

After having explained why it is useful to include control variables in empirical analyses I now introduce the specific control variables I used in my study. I implemented six socio-economic control variables which are frequently used in studies and consider personal information of the respondents are enumerated and further illustrated. Furthermore, I added six control variables particularly chosen for this empirical study.

Control variables – Personal information
What is your gender?
What year were you born in?
What is your country of birth?
In which country do you currently live?
What is your highest level of education?
What is your current status of occupation?

Figure 8: Control variables – personal information

In general I am assuming that including the six items above as control variables in my study is in accordance with standard procedures of various statistical investigations in social sciences. Often questionnaires start or end with the request to fill out some personal information in order to gather information about the sample characteristics. This argumentation is in line with what Peterson (2000) states in his book “Constructing effective questionnaires” when he writes that almost every questionnaire which directly addresses individuals comprises so-called socio-economic or demographic questions (Peterson, 2000, p. 83). Furthermore, he classifies the six demographic questions about gender, age, ethnicity, country of residence, education and occupation used in this study as common (ibid.). Additionally, Peterson (2000) claims two major reasons for asking demographic questions: Firstly, demographic questions give the researcher information about the study participants which in turn helps to make inferences to which extent the study results are generalizable and representative (Peterson, 2000, p. 84). Secondly, asking demographic questions allows comparing answers of study participants within different demographic groups (ibid.). Concerning the format and also the answer options of demographic questions no standard exists so far (ibid.). Conclusively, researchers are relatively free in designing demographic questions for their studies (ibid.).

Besides the ones mentioned by Peterson (2000) a specific reason for the inclusion of the control variable “occupation status” could be identified within the context of this study: It is assumed that people who have more spare time per day are more likely to engage in activities as crowdsourcing initiatives than people who work full-time. This argument is indirectly supported by an empirical study of Kaufmann et al. (2011) where the authors investigated the demographical characteristics of workers on the crowdsourcing platform Amazon Mechanical Turk. The authors found that almost sixty percent of all study participants either work part-time, are still in their education process or even unemployed (Kaufmann et al., 2011, p. 7). Conclusively, less than a half of the study participants worked full-time and engaged in crowdsourcing tasks on the platform (ibid.). From these findings it can be concluded that the amount of spare time people

have per day at least to some extent influences their engagement in crowdsourcing activities besides own job activities.

Lastly, the answer options for two of the six items which are supposed to gather personal information of study participants should be named. The answer options for gender are quite obvious either female or male. For age, country of birth and country of living study participants are free to type the appropriate information into a description field. Considering the level of education the study participant can choose among the options high school diploma, Abitur, Diploma, Bachelor, Master, PhD and none of these. For the status of occupation people can choose between the options student, retired, half-time employed, full-time (self-)employed or unemployed.

Control variables specifically relevant for this study
Did you previously participate in crowdsourcing projects?
I think that the capabilities I possess are sufficient to solve the proposed crowdsourcing task.
I hope to earn money by participating in such crowdsourcing tasks.
I would engage in such crowdsourcing tasks because I hope to be recognized by BMW.
I hope to enhance my professional reputation by engaging in such crowdsourcing tasks.

Figure 9: Control variables - specifically relevant for this study

Additionally, I decided to ask if people already previously participated in crowdsourcing projects. I did this because it is assumed that people who earlier gathered experiences with crowdsourcing tasks are more likely to engage in another crowdsourcing project. This argument is also in line with several psychological publications where it is argued that people are likely to follow previous patterns of behavior. For example, Ajzen (1991) states in his paper that several researchers proposed past behavior as an important predicting variable of later behavior (Ajzen, 1991, pp. 202-203). Also Ouellette and Wood (1998) identified past behavior as a crucial factor for predicting future behavior after having extensively reviewed studies about the explanation of people's behavior (Ouellette & Wood, 1998, p. 65).

Concerning people's perceived capabilities as a control variable is useful since some people might reject participating in a creative crowdsourcing project because they think that they do not possess the capabilities for this task. This argumentation is in line with the reflection about the concept of social cognitive theory by Cruickshank et al. (2010): In their publication the authors refer to the term self-efficacy which is defined as someone's beliefs about his or her abilities to perform a certain task (Cruickshank et al., 2010, p. 3). They further explain that the way people think about their capabilities to perform a specific

task influences whether they decide to take up this task (Cruickshank et al., 2010, pp. 3-4). Referring to Bandura (1986) one could argue that “[p]eople tend to avoid tasks and situations they believe exceed their capabilities [...]” (Bandura, 1986, p. 393). According to Dörnyei (2001) also Ryan and Deci talked about this phenomenon when introducing the concept of “amotivation” as a part of their self-determination theory (Dörnyei, 2001, p. 143). As reported by Dörnyei (2001) Ryan and Deci referred to amotivation as the lack of motivation which is not due to missing interest but rather due to the individual’s perception that he or she is helpless or incompetent when facing the task at hand (Dörnyei, 2001, p. 144). In accordance with the explanations of Bandura (1986), Cruickshank et al. (2010) and Dörnyei (2001) it can conclusively be argued that it is reasonable to include the way people think about their capabilities as a control variable. This is especially important in order to make sure that solely the effect of social identity on creative crowdsourcing engagement is measured.

In the following three additional items will be introduced which are supposed to consider economic motivations to engage in crowdsourcing tasks. Items for economic motivations are included for the following reasons: In this study people are asked whether they would be willing to participate in a crowdsourcing task initiated by BMW as a concrete company. Consequently, it can be hypothesized that besides a high social identity with BMW people could also be motivated by economic reasons. This is due to the fact that a concrete company is involved and people who answer that they are willing to engage in the proposed crowdsourcing task might hope to receive economic rewards of any kind by the company in question. Therefore, it can be argued that besides asking about socio-economic characteristics, earlier participation in crowdsourcing projects and perceived capabilities it is reasonable to include possible alternative explanations for crowdsourcing engagement which are directly related to economic motives as additional control variables.

Some authors found that immediate payment or the possible prospect to earn money after having completed a crowdsourcing task influences people’s engagement in crowdsourcing tasks. For example, Kaufmann et al. (2011) examined in an empirical study the motivation of people to contribute crowdsourcing tasks to the online platform Amazon Mechanical Turk and found immediate payoffs as one of the major motivational factors (Kaufmann et al., 2011, p. 7). Furthermore, the findings of Brabham (2008) who surveyed participants actively contributing to the platform of the creative photography community, iStockphoto, indicate that making money on the platform is the most important incentive to contribute (Brabham, 2008). Consequently, monetary rewards have to be considered as a possible

alternative explanation for crowdsourcing engagement and with that need to be incorporated in this study in the form of a control variable.

I included firm recognition as a control variable for my study for the reason that studies of other authors already revealed firm recognition as a significant motivational factor for contributions in online or user communities. For example Jeppesen and Frederiksen (2006) asked in their article what motivates people to contribute to firm-hosted user communities (Jeppesen & Frederiksen, 2006, pp. 45-46). In the course of their empirical study the authors clearly identified that people participating in firm-hosted user communities are highly motivated by firm recognition (Jeppesen & Frederiksen, 2006, pp. 55-56). Likewise, a study of Henning-Thurau et al. (2004) about the engagement of people in online platforms is pointing in the same direction by introducing the concept of self-enhancement as a motivational factor (Henning-Thurau et al., 2004, p. 43). Accordingly, if people are motivated by self-enhancement they are striving for a positive recognition by other persons or organizations (ibid.). In the course of their study Henning-Thurau et al. (2004) empirically support the theoretical assumption that people can be motivated to engage in user communities by the prospect to be recognized by others (Henning-Thurau et al., 2004, pp. 48-49). Consequently, it is possible that the hope to be recognized by influential persons or even the firm which initiated the crowdsourcing task has the potential to deliver an alternative explanation for the engagement in crowdsourcing tasks. That is why the factor firm recognition needs to be included as a control variable in my study.

In their study about an online community Wasko and Faraj (2005) empirically prove that the enhancement of people's professional reputation can be a motivational factor for them to contribute to the community (Wasko & Faraj, 2005, p. 50). Furthermore, also an empirical investigation about the motivation of community members of iStockphoto by Brabham (2008) has shown that the opportunity to earn a good reputation by contributing crowdsourcing work is an important motivational driver (Brabham, 2008). Also a study by Cheung and Lee (2012) about the motivation of consumers to contribute to online communities has shown that many people are motivated by the prospect of an enhanced reputation after having contributed content (Cheung & Lee, 2012, p. 222). Conclusively, it can happen that participants in my study agree to participate in the proposed crowdsourcing task because they are mainly motivated by the prospect of an enhanced professional reputation. Hence, enhancement of the professional reputation should be included as a control variable in this study.

Lastly, it should be mentioned that the item asking about previous participation is measured with the help of two answer options which are “yes” and “no”. The other four items are measured with a five-point Likert scale ranging from “strongly disagree” to “strongly agree”.

4.4. Steps for data analysis

Firstly, I am also going to construct a correlation table which will show the correlations of the mean values of all of the variables used in the model (Vogt & Johnson, 2011, p. 80). The method of correlation itself measures the linear relationship between variables (Field, 2005, p. 107). For example, one would suppose a correlation between two variables if both variables deviate from its respective means in a similar way (Field, 2005, p. 108). In the case of my study I include crowdsourcing engagement as the dependent variable, the three sub-scales of social identity respectively the one summarized social identity scale as the independent variable and all of the remaining variables as control variables. Although Bortz and Schuster (2010) differentiate in their chapter about correlation between different methods for assessing correlation in dependence of the type of scale the variables belong to (Bortz & Schuster, 2010, pp. 153-182) it also seems to be possible to include variables of different scales into one correlation matrix. For example, looking at the paper of Vorkinn and Riese (2001) reveals that they put gender which is measured with a nominal scale¹⁸ together with interval scaled¹⁹ variables such as age, income and attachment to impacted areas (Vorkinn & Riese, 2001, p. 259). Consequently, it also seems to be reasonable for my study to summarize variables with different scales in one correlation table. After having evaluated this correlation table it will be possible to detect noticeable correlations between two variables. Secondly, I have decided to firstly undertake a simple regression analysis to research to what degree the level of a person’s social identity influences his or her level of crowdsourcing engagement. This idea of using the method of a regression analysis for this study fits to the way how Field (2005) nicely describes in his book the essence of a regression analysis: “[I]t’s a way of predicting some kind of outcome from one or more predictor variables.” (Field, 2005, p. 143). Due to the fact that several control variables and also the three

¹⁸ If objects are measured in a nominal scale they are categorized (Bortz & Schuster, 2010, p. 13). Typical examples are nationality, family status or religion (ibid.).

¹⁹ If objects are measured in an interval scale it is possible to interpret the difference between two consecutive scale values because this difference remains constant for every scale value (Bortz & Schuster, 2010, p. 14). Common examples are scales for temperatures which indicate the temperature in Celsius or Fahrenheit (ibid.).

sub-factors of social identity should be taken into account as independent variables in this study it is appropriate to also conduct multiple regression analyses where an outcome is predicted from several variables (Field, 2005, p. 144). The overall concept of my empirical analysis is that in total four regression analyses will be pursued where crowdsourcing engagement will always be the dependent variable: Firstly, only social identity will be taken as an independent variable. Secondly, the three sub-factors of social identity will serve as independent variables. In the third model not only social identity will be included as an independent variable but also those variables which correlated with crowdsourcing engagement in the correlation analysis. Fourthly, the variables which correlated with crowdsourcing engagement as well as the three sub-factors of social identity will be chosen as independent variables.

5. Results

In the results section the following points will be covered: Firstly, a correlation analysis of all the variables included in this study will be conducted. Secondly, the sample will be tested for outliers with the help of Cook's and Mahalanobis distance. Thirdly, the assumptions for undertaking a regression analysis will be listed and tested. Fourthly, the four regression models applied in this study will be introduced and tested empirically. Fifthly, the results from all four regression analyses will be summarized and discussed.

5.1. Correlation analysis

A next step was to include the three sub-scales of social identity as well as the summarized social identity scale as the independent variable, crowdsourcing engagement as the dependent variable and all of the control variables into one correlation table. By that one will be able to determine the Pearson's correlation coefficient for each of the included variables. A correlation coefficient is a common measure which is used to make a statement about the effect size of a relationship between two variables (Field, 2005, p. 111). The values of Pearson's correlation coefficient usually lie between -1 and 1 (ibid.). A correlation coefficient of 1 indicates a perfect positive correlation between two variables which means that if one variable increases by a certain amount the other variables increases by the proportionate amount as well (ibid.). Whereas a correlation coefficient of -1 shows that two variables have a perfect negative relationship (ibid.). If Pearson's correlation coefficient is 0 it can be concluded that there is no relationship at all between both variables (ibid.). Here it should be remarked that I used a two-tailed test for the correlation

analysis. A two-tailed test is used if you cannot predict the nature of the relationship between variables (Field, 2005, p. 123). The correlation table I inserted below shows the correlation coefficient of all the included variables I mentioned above. Whereas one asterisk shows that the correlation is significant on a level of .05, two asterisks show that the correlation between the respective variables is significant on a level of .01 (Field, 2005, p. 126).

For the purpose of my study I looked at the correlation table and investigated which variables correlated significantly with my dependent variable crowdsourcing engagement. With a correlation coefficient of .444 and a significance value of .000 the factor age has the strongest positive correlation with crowdsourcing engagement. This means that younger people²⁰ are more likely to participate in crowdsourcing activities. Secondly, people's retirement status correlates negatively with crowdsourcing engagement while having a correlation coefficient of -.340 and a significance value of .000. This shows that people who do not work anymore are not very likely to engage in crowdsourcing activities. Thirdly, people's self-esteem positively correlates with crowdsourcing engagement by exhibiting a correlation coefficient of .298 and a significance level of .000. Consequently, people scoring high in the scale for self-esteem are more likely to engage in crowdsourcing activities. The capabilities people perceive to have is the fourth strongest factor which correlates with crowdsourcing engagement because it has a correlation coefficient of .291 and a significance value of .000. This shows that people who think that they have sufficient capabilities to solve the proposed crowdsourcing task are more likely to also engage in crowdsourcing initiatives than people who are unsure about their ability to solve the proposed crowdsourcing task. The fifth greatest effect on crowdsourcing engagement comes from social identity which features a correlation coefficient of .269 with a significant value of .000. We can conclude from this that people with a high social identity are more likely to engage in crowdsourcing than people with a low social identity. Sixthly, affective commitment positively correlates with crowdsourcing engagement with a correlation coefficient of .227 and a significance value of .000. Thus, people who feature a high affective commitment might also be more willing to participate in crowdsourcing initiatives. The seventh correlation is a positive one between self categorization and crowdsourcing engagement ($r = .209$, $p = .000$). This shows that people with high scores in the self categorization scale are also the kinds of persons to engage in crowdsourcing. The occupation status student is the eight factor which correlates with crowdsourcing engagement. It has a correlation coefficient of .208 and a significance value of .000 which

²⁰ This is because I measured the age level with the year people were born in. Consequently, people with a high age level are younger than people with a low age level.

shows that students are more likely to engage in crowdsourcing. Ninthly, the belief to be recognized by firms when engaging in crowdsourcing activities has an effect on crowdsourcing engagement and shows a correlation coefficient of .198 and a significance value of .001. Consequently, another reason for people to participate in crowdsourcing activities is the hope to be recognized by companies in order to improve their professional situation. The two last factors which correlate with crowdsourcing engagement are both related to the level of education of study participants: People who said that their highest level of education is a Bachelor are according to the correlation coefficient of .163 and significance level of .006 more likely to participate in crowdsourcing initiatives than people who declared that they have none of the listed education levels. This last group of people features a correlation coefficient of -.153 with a significance level of .011 which shows people without any of the listed education levels would rather not engage in crowdsourcing. Here, one could suppose that the finding that people with a Bachelor degree are more likely to participate in crowdsourcing activities is in line with the finding that younger people are rather willing to engage in crowdsourcing. Usually, these younger people also have a lower educational level than older people who might already have completed their studies. Furthermore, it is assumed that people who stated that they possess none of the listed levels of education usually have if anything a lower educational level than a high school degree. One could conclude from this that these study participants are either really young and are still in their first years of school or that they are very old and belong to these generations where people generally did not have the possibilities to earn a very high education level. The first argument that people who did not possess any of the listed education levels are really young can be invalidated because the youngest two study participants were born in the year 1995 and should therefore at least possess a high school degree. Therefore, it can be supposed that people without any of the named levels of education might be people from the older generations. Consequently, the fact that the educational degree none of these correlated negatively with crowdsourcing engagement can be ascribed to the fact that older people seem to be generally less interested in crowdsourcing than younger people. This idea was also supported by the correlation between age and crowdsourcing engagement.

These correlations show that apart from my assumption that social identity and its three sub-factors positively correlate with crowdsourcing engagement also seven other factors seem to influence people's crowdsourcing engagement. An overview of the correlations between all variables which correlated with crowdsourcing engagement can be found in figure 10. In the following only these variables which correlated with crowdsourcing engagement will be taken into consideration for the regression models.

At this point it should be discussed to which extent these control variables which were derived from the literature as possible influencing factors for crowdsourcing engagement at the end really correlated with crowdsourcing engagement. Additionally, it needs to be examined whether the reasons for the proposed relationship between the control variables and crowdsourcing engagement mentioned in the literature match the ones found with the help of the correlation analysis. The occupation status was named in the literature as an influencing factor for crowdsourcing engagement because it was assumed that people who are half-time employed, unemployed, retired or still in their education are more likely to engage in crowdsourcing simply because they have more spare time to do so. Although a correlation between people's occupation status and crowdsourcing engagement has been found it is not the same as proposed in the literature. In this study it was found that retired people are less likely to engage in crowdsourcing activities than students who are still in their education. Although it was argued with the help of the literature that people who previously participated in crowdsourcing projects are more likely to engage again in crowdsourcing initiatives no such a relationship was found in this study. Moreover, it has been shown that people's perception of their own capabilities influences their decision whether or not to participate in crowdsourcing projects. This supports the assumption stated by several authors that people's decisions to act in a certain way is generally strongly influenced by their own perception of whether they will manage this action with the capabilities they possess. Additionally, the correlation analysis explicitly confirms this idea in connection with the relationship between people's perceived capabilities and their willingness to engage in crowdsourcing. Furthermore, in the literature the prospect of earning money by engaging in crowdsourcing was supposed to have a positive influence on crowdsourcing engagement. This relationship could not be confirmed in this study. However, the proposed positive relationship between the hope to be recognized by firms and crowdsourcing engagement from the literature could be affirmed in this study. Conversely, the positive relationship between the expectation to enhance one's professional reputation and crowdsourcing engagement which was stated in the literature could not be shown with the help of the correlation analysis. In addition to the influencing factors on people's crowdsourcing engagement derived from the literature two demographic variables emerged as significant influencing factors: Firstly, it was found that younger people are more likely to engage in crowdsourcing. Secondly, the educational level Bachelor positively influences people's crowdsourcing engagement whereas people who declared to possess none of the listed levels of education are highly unlikely to engage in crowdsourcing.

	mean	Std. dev.	1	2	3	4	5	6	7	8	9	10	11	12
1 Age	1970.21	14.12	1											
2 Abitur	0.298	0.45	-.059	1										
3 None of these	0.158	0.36	-.483**	-.283**	1									
4 Student	0.118	0.32	.505**	-.069	-.159**	1								
5 Retired	0.082	0.27	-.588**	-.196**	.657**	-.110	1							
6 Crowdsourcing engagement	2.381	1.14	.444**	-.094	-.153*	.208**	-.340**	1						
7 Capabilities	2.439	1.01	.095	-.182**	-.139*	-.082	-.195**	.291**	1					
8 Firm recognition	1.95	0.96	.138*	-.015	-.090	.066	-.175**	.198**	.097	1				
9 Self-categorization	2.0576	0.82	.069	-.139*	.079	-.050	-.061	.209**	.265**	.223**	1			
10 Self-esteem	3.3264	0.94	-.158**	-.065	.145*	-.147*	-.004	.298**	.323**	.171**	.505**	1		
11 Affective commitment	2.4329	0.98	.055	-.148*	.131*	-.075	-.050	.227**	.305**	.203**	.828**	.633**	1	
12 Social identity	2.5155	0.81	.005	-.139*	.132*	-.096	0.047	.269**	.333**	.227**	.903**	.770**	.953**	1

Figure 10: Correlation table

5.2. Cook's distance and Mahalanobis distance

Due to the fact that I pursue regression analyses it is helpful to know how proper my overall regression model is. For this purpose it is critical to check whether the model fits the observed data or if it is rather biased by some cases (Field, 2005, p. 162). Outliers can generally be detected by looking at the differences between the values of the outcome predicted by the model and the values of the outcome observed in the sample (Field, 2005, p. 163). These differences are called residuals and represent the errors which exist in the model (ibid.). In the case that the data perfectly fit the model these residuals will be zero which means that all data fall on the regression line (ibid.). In order to interpret the residuals standardized residuals are used which are the residuals divided by an estimate of their standard deviation (Field, 2005, pp. 163-164). For obtaining outliers in the model it is also possible to look whether some cases influence the model in a certain direction (Field, 2005, p. 164). One measure which helps here is the so-called Cook's distance which measures the overall influence of single cases on the whole model (Field, 2005, p. 165). If Cook's distance exceeds the value of 1 it might be that one case might distort the model (ibid.). Additionally, another measure of influence, called leverage, can be used to detect the influence of certain outlier cases over the prediction made in the model (ibid.). If the leverage value turns out to be 1 then it can be concluded that the respective case has a complete influence over the prediction of the outcome (ibid.). In the case that none of the cases has an undue influence over the predicted outcome all leverage values should be close to the average value of the outcome variable (ibid.). For calculating leverage values Mahalanobis distance can be used which measures the distance of cases with the help of the means of the independent and control variables (ibid.). Basically, after having conducted this test one has to look for the cases with the largest values (ibid.). According to Field (2005) in large samples with about 500 cases values above 25 are problematic (ibid.).

In order to receive values for Cook's distance for my first regression model I inserted crowdsourcing engagement as the dependent variable and social identity as the independent variable. Secondly, crowdsourcing engagement remained as the dependent variable and the three sub-scales of social identity were included as independent variables. Thirdly crowdsourcing engagement was picked as the dependent variable again and those variables which correlated highly with crowdsourcing engagement in the correlation analysis and the variable for social identity were inserted as independent and control variables. For the fourth regression model only the variables for social identity was substituted by the

three sub-factors of social identity while all the other variables included in the third model also remained in the fourth model. As can be seen from the four figures below the value for Cook's distance never exceeds 1. According to Field (2005) this means that none of the included cases distorts any of the four regression models in a certain way (Field, 2005, p. 165).

According to Bortz (2010) for Mahalanobis distance only variables which are interval scaled can be used (Bortz & Schuster, 2010, p. 456). Therefore, for the third and fourth regression model only interval scaled variables were. These variables are age, perceived capabilities and firm recognition. Consequently, the dummy coded categorical variables Bachelor, none of these, student and retired were removed from the third and fourth regression model. In accordance with Stevens (2009) the maximum values of Mahalanobis distance for all four regression models are below the critical value (Stevens, 2009, p. 108). Conclusively, also the examination of Mahalanobis distance has shown that no outliers could be detected in the data at hand.

Regression model I	Minimum	Maximum	Mean	Std. Dev.	N
Cook's distance	.000	.025	.003	.004	278
Mahalanobis distance	.000	3.897	.996	.676	278

Regression model II	Minimum	Maximum	Mean	Std. Dev.	N
Cook's distance	.000	.051	.004	.006	278
Mahalanobis distance	.193	19.628	2.989	2.815	278

Regression model III	Minimum	Maximum	Mean	Std. Dev.	N
Cook's distance	.000	.078	.004	.007	278
Mahalanobi's distance	.534	15.812	3.986	2.218	278

Regression model IV	Minimum	Maximum	Mean	Std. Dev.	N
Cook's distance	.000	.052	.004	.007	278
Mahalanobi's distance	.660	20.818	6.978	3.620	278

Figure 11: Cook's and Mahalanobis distance - regression models I-IV

5.3. Four regression models

In order to be able to generalize the findings after having pursued the regression analyses it is critical that the underlying assumptions for conducting a regression analysis are met (Field, 2005, p. 169). Field (2005) names nine assumptions which have to be fulfilled in order to undertake a regression analysis (ibid.). These nine assumptions are fulfilled in this study and therefore the method of regression analysis can be used. The detailed tests for each of these assumptions can be found in the appendix.

The final step for analyzing my data is to compute the actual regression analyses. The question which has to be asked is to what degree social identity and its three sub-factors predicts people's crowdsourcing engagement even if additional variables are added to the model. In order to answer this question I pursue four steps: The first step is to take only the factor of social identity as an independent variable into account and to look how much of the variance in people's crowdsourcing engagement can be explained by their social identity. This means that in the first regression analysis only the variables social identity as the independent variable and crowdsourcing engagement as the dependent variable are included. Secondly, it will be tested whether the three sub-factors of social identity (independent variables) explain crowdsourcing engagement (dependent variable) differently in comparison to the overarching concept of social identity which was tested in the first model. The third step will be to also include all the control variables which earlier correlated with crowdsourcing engagement and to examine how the influence of social identity on crowdsourcing engagement changes if more variables are taken into account. Fourthly, the same procedure as in the third model will be done but instead of social identity the three sub-factors of social identity will be examined. An overview of the results of all the four regression analyses is given below in the form of a summarizing table:

	1				2				3				4			
	B	s.e.	Beta	p	B	s.e.	Beta	p	B	s.e.	Beta	p	B	s.e.	Beta	p
Social identity	0.378 **	0.081	.269	.000					0.250 *	0.081	.178	.002				
Self categorization					0.117	0.142	.085	.410					0.058	0.120	.042	.629
Self-esteem					0.315 *	0.090	.260	.001					0.443 **	0.081	.366	.000
Affective commitment					-0.009	0.133	-.008	.944					-0.163	0.115	-.141	.156
Age									0.030 **	0.006	.375	.000	0.038 **	0.006	.471	.000
Bachelor (education)									-0.204	0.225	-.63	.366	-0.146	0.218	-.045	.503
None of these (education)									0.440	0.219	.140	.045	0.448	0.211	.143	.035
Student (occupation)									0.332	0.259	.094	.202	0.226	0.251	.064	.368
Retired (occupation)									-0.635	0.319	-.153	.048	-0.487	0.308	-.117	.116
Perceived capabilities									0.206 *	0.062	.182	.001	0.172 *	0.060	.152	.005
Firm recognition									0.086	0.063	.072	.170	0.079	0.060	.066	.194
R-square		.073				.094				.321				.377		
N = 278; * p < .005; ** p < .001																

Figure 12: Regression analyses - regression models I-IV

The regression coefficient represents the gradient of the regression line and with that indicates how the outcome changes if the independent variable changes (Field, 2005, pp. 155-156). For social identity the regression coefficient is .378. Therefore, it can be said that if we suppose that social identity increases by one unit then crowdsourcing engagement will increase by .378 units. Furthermore, it can be said that if the regression coefficient is bigger than zero and additionally big in relation to the standard error the independent variable has a significant impact on the researcher's ability to predict the outcome (Field, 2005, p. 156). In this case both conditions are given because the regression coefficient is with a value of .378 different from zero and at the same time relatively bigger in comparison with the standard error which is .081. Consequently, it can be concluded that the independent variable social identity helps to predict people's crowdsourcing engagement. Coming back once more to the standard error it can additionally be said that it shows to what degree the value for the regression coefficient would vary across different samples (Field, 2005, p. 192). Furthermore, the standard error is used to demonstrate if the value for the regression coefficient significantly differs from zero (ibid.). One can say that if the standard error is small in relation to the value of the regression coefficient most samples are likely to have a similar value for the regression coefficient (Field, 2005, p. 151). In the case of this regression analysis it can be said that the standard error of .081 for the regression coefficient is relatively small. As a third value which can be found in the coefficients table the standardized coefficient beta comes into play. This value provides the number of standard deviations the outcome will change if the independent variable changes in one standard deviation (Field, 2005, p. 193). Here this would mean that if social identity changes in one standard deviation crowdsourcing engagement would change in .269 standard deviations. Additionally, p represents the significance of the results obtained in this first regression model (Field, 2005, p. 156). As long as this significance value is less than .05 it is generally assumed that the associated variables make a significant contribution in predicting the outcome (ibid.). In this case the p value of .000 shows that social identity significantly influences people's crowdsourcing engagement if it is the only independent variable included in the model. Lastly, the meaning and value of the correlation coefficient squared should be clarified for this model: The correlation coefficient squared represents the amount of variance in the outcome (crowdsourcing engagement) which can be explained by the model (social identity → crowdsourcing engagement) (Field, 2005, p. 148). If the value of the squared correlation coefficient is multiplied by hundred one obtains the exact percentage value which says how much percent of the variance in the outcome is explained by the model (ibid.). Consequently, 7.3 percent of the variance in people's crowdsourcing engagement can be explained by their social identity if only social identity is taken into account as an independent variable.

In the second regression model the three sub-factors of social identity were taken as independent variables and their influence on crowdsourcing engagement as the dependent variable was tested. As can be conducted from a short glance at the summarizing table of the regression results in the second model only self-esteem significantly influences crowdsourcing engagement. This can be said because its regression coefficient (0.315) is large in relation to the standard error (0.090) and has an associated significant value of .001. The two other regression coefficients are not only relatively small compared to their standard errors but are also not significant. Conclusively, only a statement about the influence of self-esteem on crowdsourcing engagement can be made: If self-esteem increases by one unit crowdsourcing engagement will increase by 0.315 units. Additionally, due to the small standard error of 0.090 in relation to the regression coefficient it is supposed that most samples are likely to have a similar regression coefficient for self-esteem. Furthermore, the standardized coefficient beta for self-esteem shows that if self-esteem changed in one standard deviation crowdsourcing engagement would change in .260 standard deviations. Moreover, the second model explains 9.4 percent of the variance in people's crowdsourcing engagement because it exhibits a correlation coefficient squared of .094.

Thirdly, a regression analysis containing social identity and all other variables which correlated earlier with crowdsourcing engagement as independent variables and crowdsourcing engagement as the dependent variable is discussed. Looking at the value of the regression coefficient of social identity for the third regression model it can be said that if social identity increases by one unit crowdsourcing engagement increases by .250 units. Due to the fact that the value of the regression coefficient b is bigger than zero and also relatively big in comparison to the standard error of .081 it can be said that social identity still has a significant impact on the prediction of the outcome in the form of crowdsourcing engagement even if additional variables are included. Additionally, the relatively small standard error in comparison to the regression coefficient b allows the conclusion that the value of social identity's regression coefficient would be very similar to this one in most other samples. Furthermore, the standardized coefficient beta shows that if social identity changed in one standard deviation crowdsourcing engagement would change in .178 standard deviations. Finally, from the significance value (.002) it can be concluded that the factor social identity still predicts the outcome in the form of crowdsourcing engagement to a considerable proportion even though other control variables are included. Furthermore, it can be said that the strength of the relationship between social identity and crowdsourcing engagement does not diminish a lot in the third model where all the correlated control variables are included. This can be seen in the difference of the regression coefficient of social identity between both regression analyses: $.378 - .250 = .128$.

Furthermore, both regression models show that social identity influences crowdsourcing engagement in a positive way. For the third model the value for the correlation coefficient squared is .321 which shows that social identity together with those independent variables which correlated earlier with crowdsourcing engagement explains 32.1 percent of the variance in crowdsourcing engagement. However, it has to be stated that apart from social identity only the factors age and perceived capabilities exert a significant influence on crowdsourcing engagement. All of the other variables which correlated with crowdsourcing engagement in the correlation analysis do not possess a significant regression coefficient in this third regression analysis. Furthermore, it should be mentioned that if we look at the values of the respective regression coefficients and the significances among these three variables social identity has the largest impact on crowdsourcing engagement.

In the fourth regression analysis the influence of the three sub-factors of social identity together with all the variables which correlated earlier with crowdsourcing engagement on the dependent variable crowdsourcing engagement was observed. From the social identity perspective it can firstly be stated that as in the second regression model also in this model only self-esteem as a sub-factor of social identity significantly influences crowdsourcing engagement. This can be said because the regression coefficient for self-esteem (0.443) is highly significant ($p = .000$) and also relatively bigger than the associated standard error (0.081). The small standard error in relation to the regression coefficient does not only indicate the fact that when self-esteem changes in one unit crowdsourcing engagement will change in 0.443 units, but it also tells us that the regression coefficient for self-esteem is likely to be very similar in most of other possible samples. Furthermore, the standardized coefficient beta shows that when self-esteem changes in one standard deviation crowdsourcing engagement will change in .366 standard deviations. Moreover, in comparison to the second regression model where the three sub-factors of social identity were tested for their influence on crowdsourcing engagement the impact of self-esteem even increases. This can be said because the value for the regression coefficient of self-esteem increased by 0.128 ($0.443 - 0.315$) in the fourth model. As one of the last points in the discussion of the fourth regression model it should be mentioned that the correlation coefficient squared indicates that this model 37.7 percent of the variance in crowdsourcing engagement. However, it should be noted that also in this model apart from self-esteem only the factors age and perceived capabilities possess a significant regression coefficient and therefore have some influence on crowdsourcing engagement. Nevertheless, measured against the sizes of the regression coefficients and the associated values of significance the impact of self-esteem on crowdsourcing engagement seems to be the largest one.

In the following lines a few concluding remarks about the results of the regression analyses and their implications on the research hypothesis together with its three sub-hypotheses are made: As can be seen from the discussion of regression model one and three the way how people score in the summarized social identity construct significantly influences their engagement in crowdsourcing activities. Even if additional variables as the ones which correlated earlier with crowdsourcing engagement are added in the analysis the impact of social identity in predicting people's crowdsourcing behavior does not diminish a lot. Sub-dividing the theoretical concept of social identity into three sub-factors as proposed by Piyathasanan et al. (2011) in regression models two and four reveals that only the sub-factor self-esteem exerts a huge influence on crowdsourcing engagement. Consequently, it can be concluded that the effect of social identity in model one and three may be predominantly impacted by the sub-factor self-esteem. Conclusively, only the second sub-hypothesis can be supported by the empirical results. Below in figure 13 the results of the four regression analyses are shortly summarized. The question which remains open for now is how meaningful it is to sub-divide the social identity concept into three different factors if only one of the factors carries weight in terms of its effect on people's behavior.

H1a: Greater self categorization increase engagement in the creative crowdsourcing process.	--- rejected ---
H1b: Greater group self-esteem increase engagement in the creative crowdsourcing process.	--- accepted ---
H1c: Greater affective commitment with the brand increase engagement in the creative crowdsourcing process.	--- rejected ---

Figure 13: Summary of the results from the regression analyses

6. Conclusion and discussion

It has been empirically shown in this thesis that explicitly self-esteem as a sub-factor of the social identity construct positively influences people's willingness to engage in crowdsourcing projects of the organization they derive their self-esteem from. Consequently, the second sub-proposition of the first theoretical proposition which Piyathasanan et al. (2011) stated in their paper could be empirically supported. However, this study has one major shortcoming which should be mentioned at this point: For crowdsourcing engagement as the dependent variable only one item was chosen which was scaled with the help of a five-point Likert scale and named "I would like to engage in creative crowdsourcing tasks of BMW similar to this in the future.". Thus, future work on this topic should probably include the

construction of a crowdsourcing engagement scale in order to be certain to work with a reliable dependent variable. Furthermore, due to the fact that the results have revealed self-esteem as the predominant factor in the whole social identity concept future studies should probably consider to adapt the scale for social identity in the way that only items regarding people's self-esteem are included. In this sense one could also question the appropriateness of the other two sub-scales of social identity for this study if they are not significant at all for crowdsourcing engagement. For example, one could argue that the items in the sub-scales for self categorization and affective commitment rather address people who are already part of the organization in question than people who might not have a connection to this organization. Consequently, it is worth thinking about the point whether the other two sub-scales used in this study are suitable for researching the overall motivation of a diverse group of people to engage in crowdsourcing projects of a certain organization. Another point is the fact that it is not certain that if people's willingness to engage in crowdsourcing increases simultaneously also the quality of the proposals resulting from crowdsourcing initiatives will improve. This might be a point for further research as well. The remaining questions are now: What do we do with the results of this study? How can draw conclusions from such a finding? In the following considerations I would like to broaden the view of the reader a little bit and stress that not only crowdsourcing engagement can be a positive outcome of a high self-esteem people derive from their membership in certain organizations, firms or groups. As the literature review with the discussion of related papers has shown the self-esteem people derive from their group membership in a certain organization positively influences people's general eagerness to step in for this organization and to act in favor of it. If people possess a high self-esteem in connection with a particular organization they are more likely to engage for this organization in any form. Generally, people show a high self-esteem if they are content with their membership in an organization or firm and feel comfortable as a member. Consequently, it can be said that a pleasant atmosphere in an organization positively affects people's self-esteem derived from this organization. And if people have a high self-esteem they are rather willing to do the organization a favor in terms of extra work or extraordinary engagement for this organization. Conclusively, what can be noted is the fact that if organizations manage it to make their members feel comfortable and content they will probably be rewarded for that in return.

Literature

- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179-211.
- Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of occupational psychology*, 63(1), 1-18.
- Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of management review*, 14(1), 20-39.
- Bagozzi, R. P., & Dholakia, U. M. (2006). Antecedents and purchase consequences of customer participation in small group brand communities. *International Journal of Research in Marketing*, 23(1), 45-61.
- Bandura, A. (1986). Social foundations of thought and action: A social cognitive approach. *Social Foundations of Thought and Action: A social cognitive approach*.
- Bergami, M., & Bagozzi, R. P. (2000). Self-categorization, affective commitment and group self-esteem as distinct aspects of social identity in the organization. *British Journal of Social Psychology*, 39(4), 555-577.
- Blader, S. L., & Tyler, T. R. (2009). Testing and extending the group engagement model: linkages between social identity, procedural justice, economic outcomes, and extrarole behavior. *Journal of Applied Psychology*, 94(2), 445.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human-und Sozialwissenschaftler. Lehrbuch mit Online-Materialien*: Springer DE.
- 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), 1-22.
- Brabham, D. C. (2013). Crowdsourcing
- Brown, R. (2000). Social identity theory: Past achievements, current problems and future challenges. *European Journal of Social Psychology*, 30(6), 745-778.
- Buchanan, B. (1974). Building organizational commitment: The socialization of managers in work organizations. *Administrative science quarterly*, 533-546.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56(2), 81.
- Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*: Harvard Business Press.
- Cheung, C. M. K., & Lee, M. K. O. (2012). What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decision Support Systems*, 53(1), 218-225. doi: <http://dx.doi.org/10.1016/j.dss.2012.01.015>
- Cruikshank, P., Edelman, N., & Smith, C. F. (2010). Signing an e-petition as a transition from lurking to participation.
- Culpepper, R. A. (2000). A test of revised scales for the Meyer and Allen (1991) three-component commitment construct. *Educational and Psychological Measurement*, 60(4), 604-616.
- Den Hartog, D. N., & Belschak, F. D. (2007). Personal initiative, commitment and affect at work. *Journal of Occupational and Organizational Psychology*, 80(4), 601-622.
- Dörnyei, Z. (2001). *Teaching and researching: motivation*: Pearson Education.
- Dukerich, J. M., Golden, B. R., & Shortell, S. M. (2002). Beauty is in the eye of the beholder: The impact of organizational identification, identity, and image on the cooperative behaviors of physicians. *Administrative Science Quarterly*, 47(3), 507-533.
- Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European journal of social psychology*, 29(2-3), 371-389.

- Enkel, E., Gassmann, O., & Chesbrough, H. (2009). Open R&D and open innovation: exploring the phenomenon. *R&d Management*, 39(4), 311-316.
- Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research*, 15(2), 195-219.
- Field, A. (2005). *Discovering Statistics Using SPSS* (Second Edition ed.). London, Thousand Oaks, New Delhi: Sage Publications Ltd.
- Frankfort-Nachmias, C., & Leon-Guerrero, A. (2010). *Social statistics for a diverse society*: SAGE Publications, Incorporated.
- Fullerton, G. (2003). When does commitment lead to loyalty? *Journal of service research*, 5(4), 333-344.
- Gangadharbatla, H. (2008). Facebook me: Collective self-esteem, need to belong, and internet self-efficacy as predictors of the iGeneration's attitudes toward social networking sites. *Journal of interactive advertising*, 8(2), 5-15.
- Gardner, T. M., Wright, P. M., & Moynihan, L. M. (2011). The impact of motivation, empowerment, and skill-enhancing practices on aggregate voluntary turnover: the mediating effect of collective affective commitment *Personnel psychology*, 64(2), 315-350.
- Haslam, S. A. (1997). Stereotyping and social influence: Foundations of stereotype consensus. In R. Spears (Ed.), *The Social Psychology of Stereotyping and Group Life* (pp. 119-143): Wiley.
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of interactive marketing*, 18(1), 38-52.
- Howe, J. (2006a). Crowdsourcing: A Definition. *Crowdsourcing - Why the power of the crowd is driving the future of business*, 2013
- Howe, J. (2006b). Customer-made: the site. *Crowdsourcing: Tracking the rise of the amateur* Retrieved 10.7., 2013
- Howe, J. (2006c). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Hyman, L., Lamb, J., & Bulmer, M. (2006). *The use of pre-existing survey questions: implications for data quality*. Paper presented at the Proceedings of the European Conference on Quality in Survey Statistics.
- Jaros, S. (2007). Measurement issues in the Meyer and Allen model of organizational commitment. *Journal of Organizational Behavior*, 6, 7-25.
- Jeppesen, L. B., & Frederiksen, L. (2006). Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science*, 17(1), 45-63.
- Kaufmann, N., Schulze, T., & Veit, D. (2011). *More than fun and money. Worker Motivation in Crowdsourcing – A Study on Mechanical Turk*. Paper presented at the AMCIS 2011 Proceedings.
- Knippenberg, D., & Schie, E. (2000). Foci and correlates of organizational identification. *Journal of Occupational and Organizational Psychology*, 73(2), 137-147.
- Knippenberg, D., & Sleebos, E. (2006). Organizational identification versus organizational commitment: self-definition, social exchange, and job attitudes. *Journal of Organizational Behavior*, 27(5), 571-584.
- Kristensson, P., Gustafsson, A., & Archer, T. (2004). Harnessing the creative potential among users*. *Journal of Product Innovation Management*, 21(1), 4-14.
- Lichtenthaler, U. (2011). Open Innovation: Past Research, Current Debates, and Future Directions. *The Academy of Management Perspectives*, 25(1), 75-93.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18(3), 302-318.
- Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of organizational Behavior*, 13(2), 103-123.
- Mael, F., & Tetrick, L. E. (1992). Identifying organizational identification. *Educational and Psychological Measurement*, 52(4), 813-824.

- Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of vocational behavior*, 61(1), 20-52.
- Michelis, D., & Schildhauer, T. (2012). *Social Media Handbuch: Theorien, Methoden, Modelle und Praxis*: Nomos Verlagsges.MBH + Company.
- Moosbrugger, H., & Kelava, A. (2011). *Testtheorie und Fragebogenkonstruktion*: Springer DE.
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychological bulletin*, 124(1), 54.
- Peterson, R. A. (2000). *Constructing effective questionnaires*: Sage.
- Pierce, J. L., & Gardner, D. G. (2004). Self-Esteem Within the Work and Organizational Context: A Review of the Organization-Based Self-Esteem Literature. *Journal of Management*, 30(5), 591-622.
- Pierce, J. L., Gardner, D. G., Cummings, L. L., & Dunham, R. B. (1989). Organization-Based Self-Esteem: Construct Definition, Measurement, and Validation. *The Academy of Management Journal*, 32(3), 622-648. doi: 10.2307/256437
- Piyathasanan, B., Patterson, P., De Ruyter, K., & Mathies, C. (2011). *Social Identity and Motivation for Creative Crowdsourcing and Their Influence on Value Creation for the Firm*. Paper presented at the Australian and New Zealand Marketing Academy
- Poetz, M. K., & Schreier, M. (2012). The value of crowdsourcing: can users really compete with professionals in generating new product ideas? *Journal of Product Innovation Management*, 29(2), 245-256.
- Riketta, M. (2005). Organizational identification: A meta-analysis. *Journal of Vocational Behavior*, 66(2), 358-384.
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics*(1), 93-107.
- Stevens, J. (2009). *Applied multivariate statistics for the social sciences*: Taylor & Francis US.
- Surowiecki, J. (2005). *The wisdom of crowds*: Random House Digital, Inc.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London: Academic Press.
- Trepte, S. (2006). Social identity theory. *Psychology of entertainment*, 255-271.
- Vogt, W. P., & Johnson, R. B. (2011). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences*: SAGE Publications, Incorporated.
- Vorkinn, M., & Riese, H. (2001). Environmental Concern in a Local Context The Significance of Place Attachment. *Environment and Behavior*, 33(2), 249-263.
- Wang, K., & Tai, C.-F. (2011, 25-27 May 2011). *The Influence of Social Presence on Continual Participation in Online Communities: The Relational View Based on Social Identity Theory*. Paper presented at the Service Sciences (IJCSS), 2011 International Joint Conference on.
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS quarterly*, 35-57.
- Wetzels, M., De Ruyter, K., & Van Birgelen, M. (1998). Marketing service relationships: the role of commitment. *Journal of business & industrial marketing*, 13(4/5), 406-423.
- Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, 5(1).
- Woisetschlager, D. M., Hartleb, V., & Blut, M. (2008). How to make brand communities work: antecedents and consequences of consumer participation. *Journal of Relationship Marketing*, 7(3), 237-256.
- Zaglia, M. E. (2013). Brand communities embedded in social networks. *Journal of Business Research*, 66(2), 216-223. doi: <http://dx.doi.org/10.1016/j.jbusres.2012.07.015>

Appendix

SPSS syntax

***** STEP 1 *****

***** recode items *****

```
RECODE Gender ("Female"=2) ("Male"=1) INTO Gender1.  
VARIABLE LABELS Gender1 'Gender numerisch'.  
EXECUTE.
```

```
RECODE Countryborn ("Estonia"=1) ("Switzerland"=2) ("Germany"=3) ("Spain"=4) ("Lithuania"=5) ("Finland"=6)  
("Austria"=7) INTO Countryborn1.  
VARIABLE LABELS Countryborn1 'Countryborn numerisch'.  
EXECUTE.
```

```
RECODE Countryliving ("Estonia"=1) ("Switzerland"=2) INTO Countryliving1.  
VARIABLE LABELS Countryliving1 'Countryliving numerisch'.  
EXECUTE.
```

```
RECODE Educationaldegree ("High school diploma"=1) ("Abitur"=2) ("Diploma"=3) ("Bachelor"=4) ("Master"=5)  
("PhD"=6) ("None of these"=7) INTO Educationaldegree1.  
VARIABLE LABELS Educationaldegree1 'Educationaldegree numerisch'.  
EXECUTE.
```

```
RECODE Occupationstatus ("Student"=1) ("Retired"=2) ("Half-time employed"=3) ("Full-time (self-)employed"=4)  
("Unemployed"=5) INTO Occupationstatus1.  
VARIABLE LABELS Occupationstatus1 'Occupationstatus numerisch'.  
EXECUTE.
```

```
RECODE Previousparticipation ("Yes"=1) ("No"=2) INTO Previousparticipation1.  
VARIABLE LABELS Previousparticipation1 'Previousparticipation numerisch'.  
EXECUTE.
```

******* recode reversely pooled items *******

RECODE Selfest1 (1=5) (2=4) (3=3) (4=2) (5=1) INTO Selfest1R.
EXECUTE.

RECODE Selfest3 (1=5) (2=4) (3=3) (4=2) (5=1) INTO Selfest3R.
EXECUTE.

RECODE Selfest4 (1=5) (2=4) (3=3) (4=2) (5=1) INTO Selfest4R.
EXECUTE.

RECODE Affectcom4 (1=5) (2=4) (3=3) (4=2) (5=1) INTO Affectcom4R.
EXECUTE.

RECODE Affectcom6 (1=5) (2=4) (3=3) (4=2) (5=1) INTO Affectcom6R.
EXECUTE.

******* create dummy variables for all categorical variables *******

RECODE Countryborn1 (1=1) (ELSE=0) INTO CountrybornD1.
VARIABLE LABELS CountrybornD1 'Estonia vs. other'.
EXECUTE.

RECODE Countryborn1 (2=1) (ELSE=0) INTO CountrybornD2.
VARIABLE LABELS CountrybornD2 'Switzerland vs. other'.
EXECUTE.

RECODE Countryborn1 (3=1) (ELSE=0) INTO CountrybornD3.
VARIABLE LABELS CountrybornD3 'Germany vs. other'.
EXECUTE.

RECODE Countryborn1 (4=1) (ELSE=0) INTO CountrybornD4.
VARIABLE LABELS CountrybornD4 'Spain vs. other'.
EXECUTE.

RECODE Countryborn1 (5=1) (ELSE=0) INTO CountrybornD5.

```
VARIABLE LABELS CountrybornD5 'Lithuania vs. other'.  
EXECUTE.
```

```
RECODE Countryborn1 (6=1) (ELSE=0) INTO CountrybornD6.  
VARIABLE LABELS CountrybornD6 'Austria vs. other'.  
EXECUTE.
```

```
RECODE Countryborn1 (7=1) (ELSE=0) INTO CountrybornD7.  
VARIABLE LABELS CountrybornD7 'Finland vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (1=1) (ELSE=0) INTO EducationaldegreeD1.  
VARIABLE LABELS EducationaldegreeD1 'High school vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (2=1) (ELSE=0) INTO EducationaldegreeD2.  
VARIABLE LABELS EducationaldegreeD2 'Abitur vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (3=1) (ELSE=0) INTO EducationaldegreeD3.  
VARIABLE LABELS EducationaldegreeD3 'Diploma vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (4=1) (ELSE=0) INTO EducationaldegreeD4.  
VARIABLE LABELS EducationaldegreeD4 'Bachelor vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (5=1) (ELSE=0) INTO EducationaldegreeD5.  
VARIABLE LABELS EducationaldegreeD5 'Master vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (6=1) (ELSE=0) INTO EducationaldegreeD6.  
VARIABLE LABELS EducationaldegreeD6 'PhD vs. other'.  
EXECUTE.
```

```
RECODE Educationaldegree1 (7=1) (ELSE=0) INTO EducationaldegreeD7.  
VARIABLE LABELS EducationaldegreeD7 'None of these vs. other'.  
EXECUTE.
```

```
RECODE Occupationstatus1 (1=1) (ELSE=0) INTO OccupationstatusD1.  
VARIABLE LABELS OccupationstatusD1 'Student vs. other'.  
EXECUTE.
```

```
RECODE Occupationstatus1 (2=1) (ELSE=0) INTO OccupationstatusD2.  
VARIABLE LABELS OccupationstatusD2 'Retired vs. other'.  
EXECUTE.
```

```
RECODE Occupationstatus1 (3=1) (ELSE=0) INTO OccupationstatusD3.  
VARIABLE LABELS OccupationstatusD3 'Half-time vs. other'.  
EXECUTE.
```

```
RECODE Occupationstatus1 (4=1) (ELSE=0) INTO OccupationstatusD4.  
VARIABLE LABELS OccupationstatusD4 'Full-time vs. other'.  
EXECUTE.
```

```
RECODE Occupationstatus1 (5=1) (ELSE=0) INTO OccupationstatusD5.  
VARIABLE LABELS OccupationstatusD5 'Unemployed vs. other'.  
EXECUTE.
```

******* STEP 2 *******

******* value labels *******

```
VALUE LABELS Gender 1 "male"  
2 "female".
```

```
VALUE LABELS Gender1 1 "male"  
2 "female".
```

```
VALUE LABELS Selfcat1 TO Capabilities 1 "strongly disagree"
```

2 "disagree"
3 "neutral"
4 "agree"
5 "strongly agree".

VALUE LABELS Previousparticipation 1 "yes"
2 "no".

VALUE LABELS Earnmoney TO Professionalreputation 1 "strongly disagree"
2 "disagree"
3 "neutral"
4 "agree"
5 "strongly agree".

VALUE LABELS Countryborn 1 "Estonia"
2 "Switzerland"
3 "Germany"
4 "Spain"
5 "Lithuania"
6 "Austria"
7 "Finland".

VALUE LABELS Countryliving 1 "Estonia"
2 "Switzerland".

VALUE LABELS Educationaldegree 1 "High school diploma"
2 "Abitur"
3 "Diploma"
4 "Bachelor"
5 "Master"
6 "PhD"
7 "None of these".

VALUE LABELS Occupationstatus 1 "Student"
2 "Retired"

- 3 "Half-time employed"
- 4 "Full-time (self-)employed"
- 5 "Unemployed".

- VALUE LABELS Countryborn1 1 "Estonia"
- 2 "Switzerland"
 - 3 "Germany"
 - 4 "Spain"
 - 5 "Lithuania"
 - 6 "Austria".

- VALUE LABELS Countryliving1 1 "Estonia"
- 2 "Switzerland".

- VALUE LABELS Educationaldegree1 1 "High school diploma"
- 2 "Abitur"
 - 3 "Diploma"
 - 4 "Bachelor"
 - 5 "Master"
 - 6 "PhD"
 - 7 "None of these".

- VALUE LABELS Occupationstatus1 1 "Student"
- 2 "Retired"
 - 3 "Half-time employed"
 - 4 "Full-time (self-)employed"
 - 5 "Unemployed".

- VALUE LABELS Previousparticipation1 1 "yes"
- 2 "no".

***** **STEP 3** *****

***** **reliability of the scales** *****

RELIABILITY

/VARIABLES=Selfcat1 Selfcat2 Selfcat3 Selfcat4 Selfcat5 Selfcat6

/SCALE ('self-categorization') ALL

/MODEL=ALPHA

/STATISTICS=DESCRIPTIVE SCALE

/SUMMARY=TOTAL.

RELIABILITY

/VARIABLES=Selfest1R Selfest2 Selfest3R Selfest4R

/SCALE ('self-esteem') ALL

/MODEL=ALPHA

/STATISTICS=DESCRIPTIVE SCALE

/SUMMARY=TOTAL.

RELIABILITY

/VARIABLES=Affectcom1 Affectcom2 Affectcom3 Affectcom4R Affectcom5 Affectcom6R

/SCALE ('affective commitment') ALL

/MODEL=ALPHA

/STATISTICS=DESCRIPTIVE SCALE

/SUMMARY=TOTAL.

RELIABILITY

/VARIABLES=Selfcat1 Selfcat2 Selfcat3 Selfcat4 Selfcat5 Selfcat6 Selfest1R Selfest2 Selfest3R Selfest4R Affectcom1
Affectcom2 Affectcom3 Affectcom4R Affectcom5 Affectcom6R

/SCALE ('social identity') ALL

/MODEL=ALPHA

/STATISTICS=DESCRIPTIVE SCALE

/SUMMARY=TOTAL.

***** means of the scales *****

COMPUTE Selfcategorization=MEAN(Selfcat1, Selfcat2, Selfcat3, Selfcat4, Selfcat5, Selfcat6).

EXECUTE.

COMPUTE Selfesteem=MEAN(Selfest1R, Selfest2, Selfest3R, Selfest4R).

EXECUTE.

COMPUTE Affectivecommitment=MEAN(Affectcom1, Affectcom2, Affectcom3, Affectcom4R, Affectcom5, Affectcom6R).

EXECUTE.

COMPUTE Socialidentity=MEAN(Selfcat1, Selfcat2, Selfcat3, Selfcat4, Selfcat5, Selfcat6, Selfest1R, Selfest2, Selfest3R, Selfest4R, Affectcom1, Affectcom2, Affectcom3, Affectcom4R, Affectcom5, Affectcom6R).

EXECUTE.

******* correlations *******

******* correlation with one social identity scale *******

CORRELATIONS

/VARIABLES=Gender1 Age CountrybornD1 CountrybornD2 CountrybornD3 CountrybornD4 CountrybornD5
CountrybornD6 CountrybornD7 Countryliving1

EducationaldegreeD1 EducationaldegreeD2 EducationaldegreeD3 EducationaldegreeD4 EducationaldegreeD5
EducationaldegreeD6 EducationaldegreeD7

OccupationstatusD1 OccupationstatusD2 OccupationstatusD3 OccupationstatusD4 OccupationstatusD5

Crowdsourcingengagement Capabilities Previousparticipation1 Earnmoney Firmrecognition

Professionalreputation Selfcategorization Selfesteem Affectivecommitment Socialidentity

/PRINT=TWOTAIL NOSIG

/STATISTICS DESCRIPTIVES

/MISSING=PAIRWISE.

******* regression analysis *******

******* regression analysis social identity scale – regression model I*******

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Crowdsourcingengagement

/METHOD=ENTER Socialidentity.

******* regression analysis sub-factors of social identity scale – regression model II*******

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Crowdsourcingengagement

/METHOD=ENTER Selfcategorization Selfesteem Affectivecommitment.

******* regression analysis social identity scale and correlated variables - regression model III*******

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Crowdsourcingengagement

/METHOD=ENTER Socialidentity Age EducationaldegreeD4 EducationaldegreeD7 OccupationstatusD1
OccupationstatusD2 Capabilities Firmrecognition.

******* regression analysis sub-factors of social identity scale and correlated variables - regression model
IV*******

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Crowdsourcingengagement

/METHOD=ENTER Selfcategorization Selfesteem Affectivecommitment Age EducationaldegreeD4
EducationaldegreeD7 OccupationstatusD1 OccupationstatusD2 Capabilities Firmrecognition.

Assumptions for a regression analysis

Variable types

The first assumption states that all independent and control variables must either be scaled at an interval level or be categorical with two categories (Field, 2005, p. 169). Furthermore, the outcome variable must be measured at an interval level and must be continuous and unbounded (ibid.). The last requirement for the outcome variable means nothing more than that there should be no constraints on the variability of the outcome (ibid.). For instance, the outcome of an outcome variable is constrained if the variable is measured with the help of a scale ranging from 1 to 10, but the gathered data only range from 3 to 7 (ibid.).

All the variables used in this study are either measured with an interval scale or are categorical variables with only two categories. Furthermore, all the outcomes of the outcome variable crowdsourcing engagement ranges from 1 to 5. Considering the fact that crowdsourcing engagement was measured with a five-point Likert scale it can be said that this variable is not constrained and continuous. Consequently, the assumption concerning the types of the variables has been met.

Non-zero variance

The non-zero variance assumption implies that the variances of the independent and control variables should not be zero (Field, 2005, p. 169). Due to the fact that according to Bortz (2010) only the variances of interval scaled variables can be calculated (Bortz & Schuster, 2010, p. 30) I only included the variables age, perceived capabilities, firm recognition, self categorization, self-esteem, affective commitment and social identity. The other variables Bachelor, none of these, student and retired which also correlated with crowdsourcing engagement and should therefore be included in the regression analyses are not listed in the table of variances.

	N	Variance
Age	278	199.537
Perceived capabilities	278	1.027
Firm recognition	278	.922
Self categorization	278	.682
Self-esteem	278	.898
Affective commitment	278	.977
Social identity	278	.665

Examining the table above shows that none of the variables which are used as an independent or control variable in the course of the regression analyses has a variance of zero. Conclusively, also the non-zero variance assumption is satisfied.

No perfect multicollinearity

This assumption states that there should be no perfect linear relationship between two or more of the control or independent variables (Field, 2005, p. 170). As a consequence these variables should not correlate too highly (ibid.).

A first step for identifying multicollinearity is checking whether any of the independent or control variables in the correlation matrix constructed earlier correlates very highly with another independent or control variable (Field, 2005, p. 175). A high correlation would be present if Pearson's correlation coefficient lies above .80 or .90 (ibid.). Observing the correlation table it can be concluded that self categorization, affective commitment and social identity correlate highly with each other. Although the assumption explicitly states that no control or independent variable should correlate too highly with another of these variables I think that it is only naturally that two sub-factors of social identity correlate with each other and also with social identity itself. This is because as the review of the literature has shown all of the three sub-factors of social identity belong to the overarching concept of social identity and should therefore be close to each other in terms of correlation. Furthermore, it can be assumed that at least the high correlation between the sub-factors of social identity and the overarching concept of social identity itself is not that problematic for my study design. This is because in none of my four regression models the three sub-factors and social identity will be considered simultaneously – it is always the case that either the three sub-factors are included as independent variables or the overarching concept of social identity is examined as an independent variable. Interestingly, the correlation coefficients between self-esteem

and the other two sub-factors and social identity only range between .505 and .770. Consequently, it can be said that apart from two of the sub-factors of social identity and social identity itself all of the other independent and control variables (age, perceived capabilities, firm recognition, Bachelor, none of these, student and retired) fulfill the assumption that no perfect multicollinearity should exist between them.

These findings can be further explored by screening the collinearity statistics for these control and independent variables (Field, 2005, p. 260). In this table one can observe the eigenvalue, condition index and variance proportion for each of these variables (ibid.). Usually, one says that if any of the eigenvalues in this table are much larger than the other eigenvalues the regression model is not very accurate and can be affected by small changes in the independent respectively control variables or outcome variable (ibid.). The condition indexes are another expression of the eigenvalues and can be calculated by taking the square root of the ratio of the largest eigenvalue and the eigenvalue from which the condition index should be calculated (Field, 2005, pp. 160-161). For my data at hand the condition index for the last dimension is considerably larger than the previous ones. According to Field (2005) this already indicates a probable multicollinearity problem (Field, 2005, p. 161). The next part of the collinearity statistics table includes the variance proportions. Generally, a variance of a variable can be split into the eigenvalue and the variance proportion (ibid.). The value of the variance proportion shows how much of the variance in a regression coefficient of an independent or control variable can be ascribed to the eigenvalue. In order to detect multicollinearity one needs to look at control and independent variables with high variance proportions on the same small eigenvalue (ibid.). In the table it can be seen that self categorization (.88), self-esteem (.71), affective commitment (.89) and social identity (.97) all possess very high variance proportions on the same small eigenvalue (.003). Consequently, these variables correlate highly with each other. This finding is also supported by the correlation analysis discussed above. The variance proportion values and eigenvalues of the remaining variables appear to be normal. Conclusively, this short observation of collinearity diagnostics has supported the findings I gained by examining the correlation table. Here again the fact that the values of the three sub-factors of social identity and social identity are suspicious in terms of their multicollinearity is only natural because all variables are related to the same overarching concept of social identity.

Dimension	Eigen- value	Condition Index	Variance proportions											
			Constant	Age	Perceived capabilities	Firm recognition	Bachelor	None of these	Student	Retired	Self categorization	Self- esteem	Affective commitment	Social identity
1	7.132	1.000	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
2	1.705	2.045	.00	.00	.00	.00	.06	.07	.06	.06	.00	.00	.00	.00
3	1.156	2.484	.00	.00	.00	.00	.08	.06	.09	.09	.00	.00	.00	.00
4	.300	4.878	.00	.00	.00	.00	.41	.31	.42	.24	.00	.00	.00	.00
5	.298	4.890	.00	.00	.00	.00	.32	.46	.20	.32	.00	.00	.00	.00
6	.165	6.579	.00	.00	.09	.85	.01	.00	.03	.00	.00	.00	.01	.00
7	.115	7.861	.00	.00	.32	.00	.08	.07	.05	.02	.04	.00	.06	.01
8	.085	9.157	.00	.00	.56	.13	.00	.00	.01	.04	.00	.00	.00	.00
9	.041	13.161	.00	.00	.01	.00	.00	.01	.01	.00	.07	.23	.01	.00
10	.003	48.138	.00	.00	.00	.00	.02	.01	.00	.00	.88	.71	.89	.97
11	1.047E- 005	823.370	1.00	1.0 0	.00	.00	.02	.01	.13	.21	.00	.06	.02	.02

Predictors are uncorrelated with 'external variables'

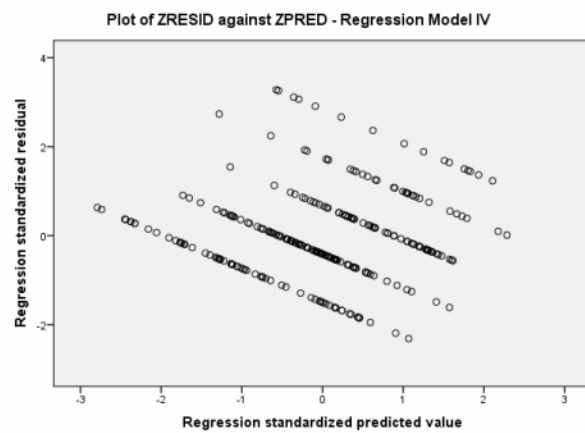
This assumption states that variables which are not included in the regression model but exert an influence on the outcome (external variables) should not correlate with any of the variables included in the regression model (Field, 2005, p. 170). In my investigation I firstly included all the variables the literature revealed as important influencing factors for crowdsourcing engagement into my correlation analysis. Later I only pursued with those variables which significantly correlated with crowdsourcing engagement in my correlation analysis. Consequently, all possible influencing factors were inherent in the beginning of my investigation in the form of variables which were inserted in the correlation analysis either as independent or control variables. That is why there are no other known external variables which could probably impact crowdsourcing engagement and which have not already been considered by me. Conclusively, in the case of my study testing for this assumption would not make any sense because of the lack of alternative external variables.

Homoscedasticity and linearity

The assumption of homoscedasticity entails that the variance of the residual terms should be constant at each level of the independent and control variables (Field, 2005, p. 170). This means that the residuals at each level of these variables should have the same variance (ibid.). Heteroscedasticity exists if these variances are very unequal (ibid.). One option to make sure that the variances possess homoscedasticity is to study the regression plots (Field, 2005, p. 181).

In order to fulfil the assumption of linearity the mean values of the outcome variable for each increment of the independent and control variables should be situated along a straight line (Field, 2005, p. 170). This assumption implies that a linear relationship is examined (ibid.).

For testing the assumptions of homoscedasticity and linearity the same plot can be used (Field, 2005, pp. 202-203). Due to the fact that the dots in all four plots for my regression models do not have the form of a funnel and also do not reveal any sort of curve it can be said that the assumptions of linearity and homoscedasticity are given in all four regression models (Field, 2005, pp. 202-203).



Independent errors

An independence of errors is given if the residual terms are uncorrelated for any two observations (Field, 2005, p. 170). This is also said to be a lack of autocorrelation and can be tested with a Durbin-Watson test (ibid.). Usually, the test statistic varies between 0 and 4 while a 2 means that the residuals are uncorrelated (ibid.). The closer the value is to 2 the more likely it is that the errors are independent (ibid.). For pursuing the Durbin-Watson test I firstly included crowdsourcing engagement as the dependent variable and social identity as the independent variable (regression model I). Secondly, crowdsourcing engagement remained the dependent variable while for the independent variable all three sub-factors of social identity were inserted (regression model II). Thirdly, as the dependent variable crowdsourcing engagement was chosen once more and as independent variables the variables which correlated earlier in the correlation analysis with crowdsourcing engagement and social identity were picked (regression

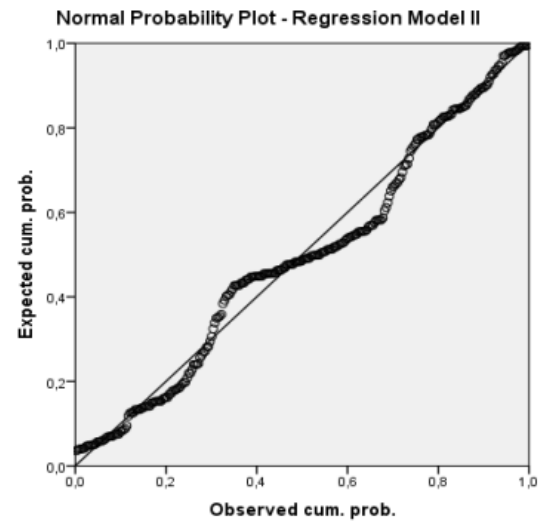
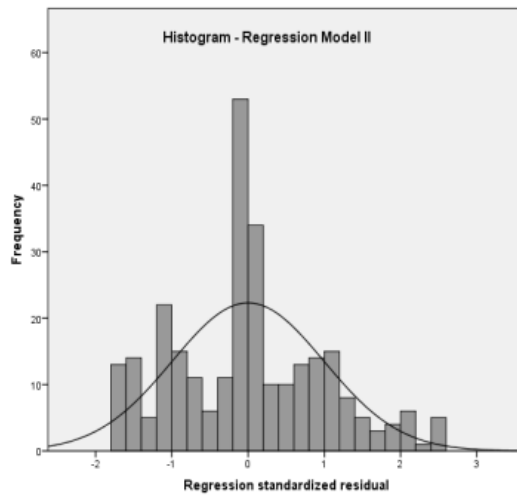
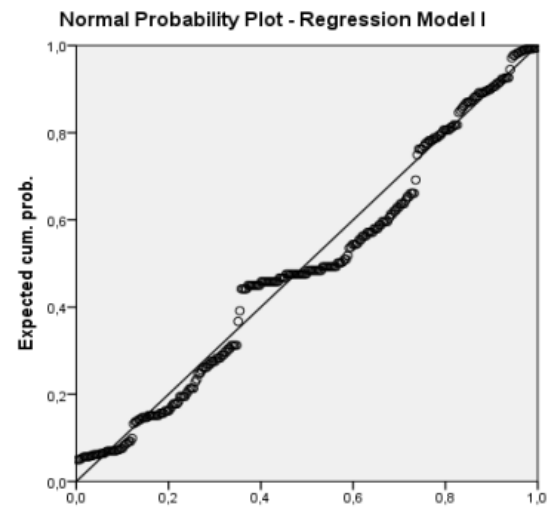
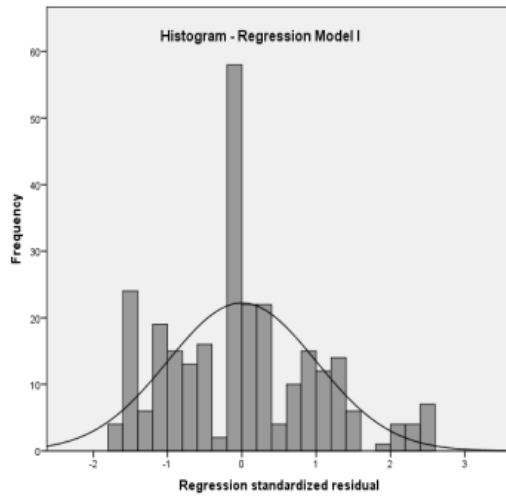
model III). Lastly, crowdsourcing engagement was taken as the dependent variable and besides the variables which correlated with crowdsourcing engagement also the three sub-factors of social identity were included as independent variables (regression model IV). As can be seen below all of the Durbin-Watson values for the four models are very close to 2 which means that an independence of errors can be assumed.

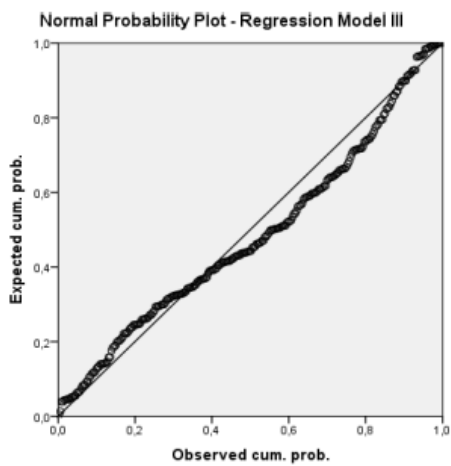
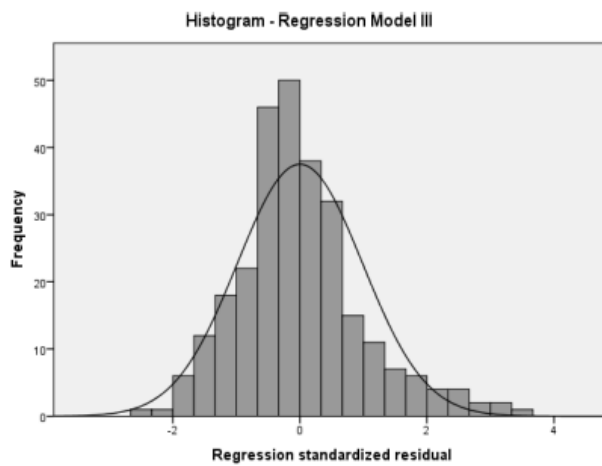
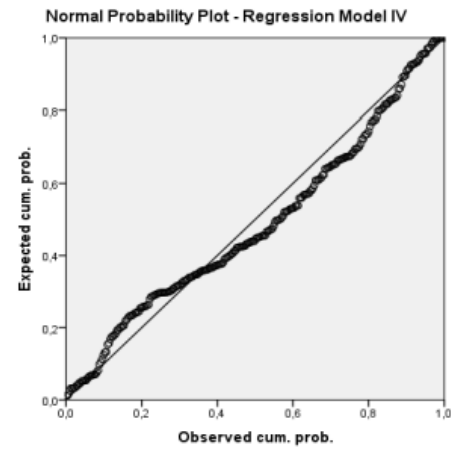
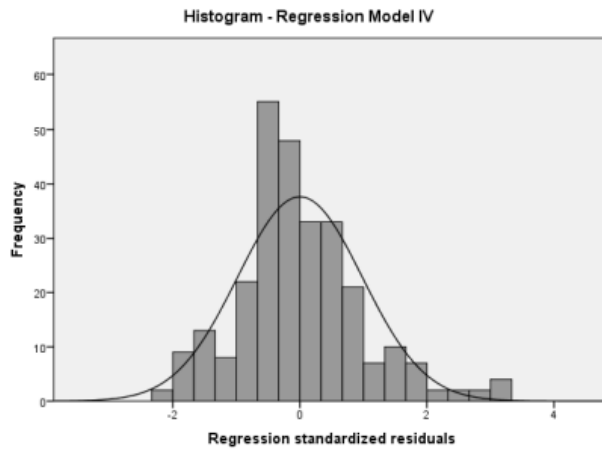
Durbin-Watson statistic

Regression model I	Regression model II	Regression model III	Regression model IV
1.945	1.952	1.878	1.917

Normally distributed errors

According to this assumption all the residuals in the model should be random, normally distributed variables and should exhibit a mean of zero (Field, 2005, p. 170). In order to test for normally distributed errors examining a histogram and a normal probability plot is helpful (Field, 2005, p. 204). If the histogram looks like a bell-shaped curve then the distribution of errors is very likely to be normal (ibid.). Furthermore, the straight line in the normal probability plot represents a perfectly normal distribution of errors (Field, 2005, pp. 204-205). Consequently, the closer the observed residuals are to the straight line the more likely it is that they are normally distributed (ibid.). For detecting normally distributed errors in my regression models I firstly inserted crowdsourcing engagement as the dependent variable and social identity as the independent variable. Secondly, I also used crowdsourcing engagement as the dependent variable and additionally introduced the three sub-factors of social identity as independent variables. For the third model crowdsourcing engagement was picked as the dependent variables once more and in addition the variables which correlated earlier with crowdsourcing engagement as well as the variable for social identity were inserted as independent variables. Fourthly, crowdsourcing engagement as the dependent variable and those variables correlating highly with crowdsourcing engagement and the three sub-factors of social identity as independent variables were chosen. As can be observed from the tables below the assumption of normally distributed errors is met for all four regression models. This can be said because all of the four histograms feature a bell-shaped curve and all of the observed residuals in all four of the normal probability plots are quite close to the straight line which represents a perfect normal distribution.





Independence

The assumption of independence states that all of the values of the outcome variable are independent which means that each value of the outcome variables stems from a separate entity (Field, 2005, p. 170). Due to the fact that I only collected data once and did not do any repeated measurements it can be assumed that the values for my outcome variable crowdsourcing engagement are independent. Additionally, all of my study participants selected an answer option for the item crowdsourcing engagement. Consequently, it can be said that an independence of values of the outcome variable is given.

Online questionnaire

Questionnaire: Crowdsourcing for BMW

Dear survey participants,

I am currently writing my master thesis about the motivation of people to participate in crowdsourcing activities. I am trying to find out how fan club membership and its connected experiences may influence people's engagement in crowdsourcing projects. That is why I am considering two groups of study participants: BMW fan club members and non-members.

Please support my research by spending ten minutes of your precious time to answer this questionnaire. This questionnaire is kept absolutely anonymous and no personal data will be saved or used for other issues than for my research. If you are interested in the results of my thesis I will be happy to forward them to you as soon as I have them.

The stars at the end of each question or statement indicate that you are kindly requested to answer these points.

Thank you for your support!

Ina Starke (inastarke@googlemail.com)

* Erforderlich

What is your gender? *

☒ Male

☐ Female

What year were you born in? *

What is your country of birth? *

In which country do you currently live? *

What is your highest level of education? *

"Sonstiges" means "something else".

- ☐ High school diploma
- ☐ Abitur (general qualification for university entrance)
- ☐ Diploma
- ☐ Bachelor
- ☐ Master
- ☐ PhD
- ☐ None of these
- ☐ Sonstiges:

What is your current status of occupation? *

- ☐ Student
- ☐ Retired
- ☐ Half-time employed
- ☐ Full-time (self-)employed
- ☐ Unemployed

Are you a member of a BMW fan club? *

- ☐ Yes
- ☐ No

When someone criticizes BMW it feels like a personal insult. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I am very interested in what others think about BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

When I talk about BMW I usually say 'we' rather than 'they'. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

BMW's successes are my successes. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

When someone praises BMW it feels like a personal compliment. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

If a story in the media criticized BMW I would feel embarrassed. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I think BMW has little to be proud of. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I feel good about BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I have little respect for BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I would rather not tell that I belong to BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I would be very happy to spend the rest of my career with BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I enjoy discussing BMW with people outside BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I really feel as if BMW's problems are my own. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I do not feel 'emotionally attached' to BMW. *

1 2 3 4 5

Load securing



Current situation - Luggage restraint systems are mandatory for commercial cars in Germany and have to consist of a separator and clamping points (a separation net is just relevant for hatch, tourings and X).

The separator has to hold 2 cubes with 18 kg in the lower part of the boot and 1 cube with 10kg near the roof. In the lower position (rear seatback positioned and material loaded behind the back) just the reference block with 10 kg over the back seats has to be held. The forward displacement of the backing facility may account max. 300mm.

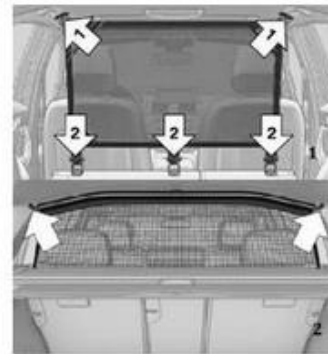
In the front position nearly the same conditions are counting, unless the adjustment of the 18kg reference block allows a collision with the backing facility.

In a boot of normal size, the lower part of the separator usually consists of the rear seat backrest, for the upper part of the boot a luggage net is used.

In an enlarged boot (by folding down the rear seat backrest), the complete separation function is done by the luggage net.

Two types of luggage nets exist:

- Stand-alone net with fastening pins for the roof and the floor panel, typically used in hatch with hard boot cover (picture 1).
- Luggage net which can be pulled out of its cartridge and fixed in the roof, typically used in station wagon with a soft boot cover (picture 2).



Disadvantages



- Handling of the luggage net (hatch), especially with enlarged boot: you have to crawl into the car and hitch the fastening pins.
- Luggage can't be restrained when only one part of the rear seat backrest is folded down.
- Storage of the luggage net, when it isn't needed.
- Handling of the heavy net cartridge (station wagon), when enlarging the boot.

Expectations for the contest: Thinking about a new separation unit (luggage net), leading to new concepts or ideally constructions.

Boot cover



Current situation

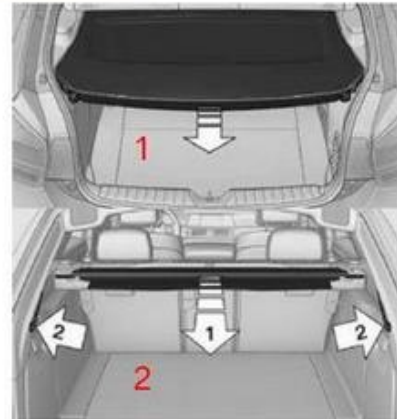
Two types of boot cover:

- Hard boot cover for hatches, GTs and SACs (SAC=X4 and X6) (picture 1).
- Soft boot cover with cartridge for station wagons (picture 2).

Challenge

- Handling and storage of the boot cover (with cartridge) and especially of the hard boot cover.
- Adaptions of the boot cover to the adjustable backrest angle, even if the angle is separately adjustable for each part of the backrest.

Expectations for the contest: We are searching for new and innovative ideas for boot coverage as well as for new concepts or ideally constructions.



I would like to engage in creative crowdsourcing tasks of BMW similar to this in the future. *

Before you choose an answer please read the following information: A crowdsourcing task is a task which a firm does not want to handle internally but rather opens up for the public to solve it. By that the firm hopes to find any person or group of people who has a clever solution to the outsourced task which can be implemented later. The pictures above will help you to imagine the creative crowdsourcing task initiated by BMW.

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I think that the capabilities I possess are sufficient to solve the proposed crowdsourcing task. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

Did you previously participate in crowdsourcing projects?

- ☐ Yes
- ☐ No

I hope to earn money by participating in such crowdsourcing tasks. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I would engage in such crowdsourcing tasks because I hope to be recognized by BMW. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

I hope to enhance my professional reputation by engaging in such crowdsourcing tasks. *

1 2 3 4 5

Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

Thank you very much for answering this questionnaire! You supported my research a lot with that! Please submit your answers by pressing "Senden". If you are further interested in my research do not hesitate to contact me via mail (inastarke@googlemail.com) or write down your email address below.

Senden

If you follow this link <https://docs.google.com/forms/d/1cGUmzQdqgOEjn42zMRpLcQyv-f68lo6YzL36LNimzks/viewform> you will be able to view the questionnaire online.