AN INQUIRY INTO THE MATHEMATICAL FORM OF THE RELATIONSHIP BETWEEN MAJORITY GROUP SIZE AND THE PROBABILITY OF FOLLOWING THE MAJORITY

THE EFFECT OF ABSOLUTE MAJORITY GROUP SIZE ON THE PROBABILITY OF FOLLOWING THE MAJORITY IN ANONYMOUS DECISION-MAKING

Charlotte Hagenaars

Abstract

The aim of this study was to provide more insight on the social influence effect of majority group size on the probability of following the majority, by looking at the mathematical form of this relationship. This was done by an analysis of experimental data showing the inclination towards choosing one out of two binary options in an anonymous decision-making situation. The results showed that the probability of following the majority depended on the absolute size of the majority group. The mathematical form that showed to be significant was a concave function, instead of a linear function. This indicated that the social influence of the majority group initially increases when this group becomes bigger, but tends to have a smaller effect when the group continues to increase in size.

Keywords

Majority influence, social influence, majority group size, popularity, concave function

Utrecht University
Bachelor Thesis Sociology

Supervisor: Prof. dr. Arnout van de Rijt
Second reader: Prof. dr. Ineke Maas

13-06-2018

Utrecht University

CONTENTS

1. INTRODUCTION

We are often faced with situations in which we quickly have to make a decision between two alternatives, without being completely sure which is best. In these brief moments of uncertainty, the choices we make are regularly dependent on the information we get from the people around us (Crutchfield, 1955). If the majority of a group gives the impression that a particular option is the right one, people are likely to imitate the behavior of their group members and conform to this perceived truth (Crutchfield, 1955; Young, 2009). Personal beliefs on the matter are ignored and herding behavior comes into place (Banerjee, 1992). This process of social influence is rooted in human behavior and changes group dynamics, leading to new norms and collective behavior (Lopez-Pintado & Watts, 2008). However, it remains unclear how big this absolute majority has to be for an individual to adjust their choice to the majority belief and to what extent the size of this majority group even matters (Bond, 2005). In order to shine a light on the dynamics of these binary decision-making situations, this research will look into the mathematical form the of the relationship between absolute majority size[1] and the probability of following the majority.

Most research on the effect of majority group size on social influence has been performed within the social psychological discipline, focusing on individual characteristics rather than group dynamics (Fischer & Vauclair, 2011). Asch (1951) found that conformity occurred when the group seemed representative and that additional agreement only confirmed initial legitimacy, leading to increasing social influence till a majority size of three people. Latané (1981) argued that increasing the majority group size leads to higher levels of conformity, but with less impact for every additional person. Tanford & Penrod's social influence model showed an S-shaped function of conformity and group size, leading to a fast initial increase of influence of the first majority group members, but less influence for additional majority group members (Tanford & Penrod, 1984). As Bond (2005) argues, the results of most studies using these different theoretical arguments are mixed and ambiguous.

Research on informational cascades has shown that in sequential decision making, social influence increases when the majority size accumulates up till the point where no new information is included in the decision-making process (Bikhchandani, Hirshleifer & Welch, 1998; Anderson & Holt, 1996). Informational cascades start when two actors make the same decision sequentially, after which the third actor will follow and ignore his own signal. Actors

[1] Absolute majority size was chosen over relative majority size, because this relationship is currently being inspected by Veldkamp (2019). Even though most studies have focussed on absolute majority group size instead of relative majority size, no concrete answers on the research question have been found so far.

making the decision later on, will assume that the third person's decision did not use their individual signal and will follow the majority, leading to no additional increase in social influence (Banerjee, 1992). The cumulative advantage model (DiPrete & Eirich, 2006) and the preferential attachment mechanism (Barabási & Albert, 1999) however, show that initial advantages in popularity, increase positive feedback, leading to linear increase in the probability of following the majority when it becomes bigger. Even though there are some similarities among these theories, finding out to what that extent absolute majority group size actually has effect on following the majority and which mathematical form explains this relationship best, will provide interesting insights in the field of social influence.

This empirical uncertainty is strengthened by the societal need for more insight in the relationship between majority group size and social influence. Even though, the average answer of groups might often be correct, leading to a certain wisdom of the crowd (Galton, 1907; Woolley et al., 2010), the social influence of a small majority giving the incorrect answer can be detrimental for this outcome (Lorenz, Schweitzer & Helbing, 2011). This often leads to the survival of misconceptions and incorrect information (Lazer et al., 2018). Shining light on this process of following the majority based on its group size can be useful in getting more insight in ways of stopping the spread of undesired, incorrect information. With the rise of fake news in the internet age, leading to the eruption of longstanding institutional frameworks fighting misinformation, insight in the spread of erroneous behavior is necessary (Lazer et al., 2018). On a brighter note, having a better understanding of this process can help improve group decision-making in small groups, such as small organizations, teams or juries where objective and correct binary decisions often have to be made. Clarity about when people will follow the majority answer can also be applied when using social information as a nudge to improve for example healthy and environmental friendly behavior (Marteau et al., 2011). When it becomes clear to what extent absolute majority size matters for following the majority, giving these nudges in the desired direction might become easier.

Due to this disagreement in and across different disciplines, additional research on the exact relationship between absolute majority size and the probability of following the majority is needed. Using the mathematical form as basic framework, will be useful for future analysis and will provide a clear answer to what extent bigger majority sizes are interesting to inspect. Not only has most research been conducted more than ten years ago, it also used relatively small samples, merely allowing to look at small majority sizes. This was mainly because most previous research focused on majority group sizes comparable to those found by Asch (1951), who found an significant effect till a majority size of three, which might not be big enough to give a proper understanding of this relationship (Bond, 2005). This research however, used a relatively big sample which was retrieved in multiple sessions, allowing to make multiple independent observations on a group level. This permitted to look at majority group sizes of twelve members, which allowed for a clear test of the mathematical form of the relationship between majority group size and the probability of following the majority, thereby giving additional information of the social influence of bigger majority sizes.

Furthermore, most previous experiments on decision-making in groups used questions which had a very obvious correct answer, in order to clearly show that the respondents that chose to ignore their personal information and follow the majority rule were influenced by this group (Bond, 2005). This research focused on more realistic decision-tasks, where respondents due to differing question difficulty were not always sure what the correct answer was. Previous research found that more difficult tasks can increase the probability of following the majority (Lucas, Firestone & Baltes, 2006). This makes this experiment applicable to more real-life decision-making processes, where tasks also differ in difficulty. Not only did the question difficulty differ, the decision-process was also more dynamic. This dynamic was created by using a cyclic order in which every subject was appointed to a different starting position. This led to the presence of different majority sizes at every choice opportunity, creating a more realistic decision process.

In order to provide more clarity on this scientific debate and to address the before mentioned societal needs, the following research question will be answered: '*To what extent does the degree of social influence depend on the size of the majority group in anonymous decision-making?*'. The theory section of this paper will provide a deeper understanding of the relationship between majority group size and the degree of social influence by focusing on several of the major theories on this topic. Based on this theoretical framework several hypotheses were derived that have been tested in order to give an appropriate answer on the research question. After that, the experimental binary decision-making design of the study, the operationalization of the variables and the methods used for the analysis will be carefully explained. The results of the analysis on the hypotheses will be shown and elaborated on in the results section. These results and implications will be discussed in the last paragraph, followed by the limitations of this research and the recommended possibilities for future research.

2. THEORETICAL FRAMEWORK

In this theoretical framework the potential mechanisms explaining the influence of majority size on conformity will be explained in detail. First of all, the general relationship between the importance of majority size and social influence will be explained, from which the first hypothesis will be derived. Secondly, several major contradicting theories on the exact shape of the relationship between majority size and the degree of social influence will be explained, leading to the remaining two contradicting hypotheses. The mathematical forms of the last hypotheses will be displayed in a figure, in order to allow for easier interpretation of the predicted form of the relationships.

2.1 Relation of majority size and the probability of following the majority

Social influence occurs when 'a person's behavior, beliefs, opinions or attitudes are transformed or controlled by some kind of social communication' (Oxford Dictionary of Psychology, 2014). Research on social influence shows that in many real-life situations, independent perceptions are ignored and influenced by those of the group around them (Cialdini & Goldstein, 2004). This process of following the group occurs in many different settings, from voting for the proposed political winner based on voting poll predictions (Cukierman, 1991) to the spread of innovations (Young, 2009) and choosing music based on popularity in artificial cultural markets (Salganik, Dodds & Watts, 2005). When making a dichotomous decision, people tend to assess the number of sources that are in favor of their position as well as the number of targets that have a preference for the alternative (MacCoun, 2012). This assessment shows them social consensus information of which alternative option is preferred by the majority, which makes people want to increase consensus by following the majority of either targets or sources (Asch, 1951). There are several micro level explanations for following the majority decision. The first is normative social influence, which occurs when a person changes his or her behavior in order to present a certain identity to the group (Deutsch & Gerard, 1955). This behavior occurs to impress group members and prevent from deviating from the group norms (Deutsch & Gerard, 1955). Informational social influence, on the other hand, occurs when a person changes his or her behavior or beliefs based on useful information retrieved from others' opinions or behavior in order to improve their personal decision (Deutsch & Gerard, 1955). Research shows that both types of social influence can occur under majority influence, even simultaneously (Insko et al., 1983). The importance of majority influence has been established a long time ago (Bond, 2005). While minorities can also influence decision and group dynamics, people tend to follow the majority rule, because of its perceived correctness (MacCoun, 2012).

While the role of majority influence has been researched extensively, the importance of the size of this majority group in anonymous decision-making remains less clear (Bond, 2005). Asch (1955) was one of the first to show the importance of majority group size. He showed that in conformity experiments, where there was a unanimous majority and no anonymity, the social influence of the majority increased until a majority size of three and remained stable after that (Asch, 1955). However, the study of Asch as well as others based on his results were focused on normative social influence, where people are quite certain that they are giving the incorrect answer when following the majority group (Bond, 2005). When studying anonymous sequential decision-making however, normative social influence is less likely to occur, because the lack of in-group and out-group identities (Deutsch & Gerard, 1955). Other theories, such as the social impact theory (Latané, 1981) and the social influence model (Tanford & Penrod, 1984), also argue that the probability of following the majority answer depends on the absolute size of the majority group and predict both normative and informational influence. The preferential attachment model (Barabási & Albert, 1999) and the cumulative advantage model show that the probability of choosing a particular option increases due to positive feedback of initially popular options, created by an increase in absolute majority size. Another major theory on following the majority based on majority size is informational cascades theory, which comes from a more economic perspective (Bikhchandani et al., 1998). This theory is based on the model of herding behavior explained by Banerjee (1992). People are seen as rational actors who can end up in situations where they have to make binary decisions subsequently (Bikchandani et al., 1998). When two previous actors have chosen to make the same choice, the third actor is unable to make a perfect assessment of the correctness of both alternatives and chooses to ignore his or her personal information and mimic the behavior or the previous actors. At this point the decision becomes uninformative to the next decision-makers and no additional social influence is wielded on the actors after him or her. Thus, during the first decisions, the group wields an increasing degree of social influence on the decision-maker through his or her perception of their accumulating knowledge. However, when this tipping point of adding no additional knowledge by extra group members has been reached, the degree of social influence of the majority group remains stable.

Even though these theories expect different types of relationships between majority size and the degree of social influence, their core assumptions all explain that the size of the majority group matters. Based on these assumptions it is expected that the proba of following the majority are dependent on the size of the majority group, leading to the following hypothesis.

Hypothesis 1: The probability of following the majority group depends on the size of the majority group

2.2 The shape of the relationship between majority size and the probability of following the majority

While most research on the importance of group size on social influence suggests that this relationship indeed exists, there is a lot of ambiguity about what this relationship exactly looks like. Therefore, two hypotheses will be derived showing the predicted mathematical forms of these contradicting theories in the following theory sections.

2.2a Llinear relationship of majority size and the probability of following the majority

Research on differences in choice popularity has focused on occasions where small initial majorities choosing a certain option can lead to increasing inequality between choices (Perc, 2014). A model explaining this phenomenon is the cumulative advantage model, which has been widely used in Sociology (DiPrete & Eirich, 2006). While it usually focuses on inequality among different societal groups, it applies to the relationship of majority size and the probability of following the majority as well. The theory is based on the Matthew effect, that showed that scientists who gained more recognition in their early careers, were more likely to have incremental success (Merton, 1988). However, recent research shows that this effect is a more general mechanisms in situations where small advantages in an early stage will grow larger over time, leading to increasing inequality (DiPrete & Eirich, 2006; Huber, 1998). This process works through “success breeds success’-dynamics, meaning that popularity of an option in early stages leads to positive feedback on the long run, leading to a higher probability of being chosen in later stages (DiPrete & Eirich, 2006). This increasing popularity occurs, because the probability of following the initial majority answer increases with every person that chooses it, showing increasing odds of following the majority for every incremental step in majority size. Cumulative advantage applies to many social influence situations, such as artificial cultural markets (Salganik, Dodds & Watts, 2005) and to a certain degree to success in awards, money and quality ratings (van de Rijt et al., 2014). Another mechanism explaining this linear impact of majority size on the probability of following the majority is the preferential attachment mechanism (Barabási & Albert, 1999). In their research they showed that in some networks a preferential attachment mechanism exists. This indicates that the more connections a node has initially, the more likely it is to receive new connections due to the automatic reinforcing mechanisms of this initial inequality (Barabási & Albert, 1999). When applying this to our research question, this would indicate that when a small majority exists, showing initial inequality in popularity among options, more people will be likely to choose this majority answer due to the positive feedback it receives. This process will accumulate, leading to an exponential growth in the odds of following the majority when the answer is initially more popular. Based on these assumptions the following hypothesis can be derived.

Hypothesis 2: Majority size has a linear impact on the probability of following the majority

2.2b Concave relationship of majority size and the probability of following the majority

Several experimental studies have pointed towards a concave, r-shaped, relationship between majority size and the degree of social influence (MacCoun, 2012). One of the major theories explaining the relation between majority size and social influence is the social impact theory (Latané, 1981). This theory has been used in describing group dynamics and polarization (Nowak, Szamrej & Latané, 1990). It combines the before mentioned informational and normative influence as attributing factors to the importance of majority size. Latané (1981), explained that in group dynamics both these effects strengthen the probability of following the majority when the number of social influence sources increases. For informational influence this occurs when people see that many others have chosen for a particular option, which strengthens their belief in the correctness of this answer, making them more likely to choose the same option. Normative influence however, works through the group’s power to punish deviations. Since bigger groups have more opportunities to punish deviant actors, social influence increases with a bigger majority. However, the impact of every additional actor becomes lower. Latané (1981), explains this based on a psychosocial law indicating that the pressure from the first person influencing you is most intruding, while later sources have less social force. This leads to a negatively accelerating curve of social influence, expressed by the following equation for less than one additional source:

$$I = sN^t$$

I = impact, *s* = a scaling constant, *N* = number of sources of influence and the exponent *t* is a value less than one.

In this equation *I* is the social impact, or social influence, that an individual feels from the actions or thoughts of at least one other person on his or her own attitudes or behavior. This impact is influenced by the number of sources that exert influence on the actor (*N*) times the scaling constant (*s*), which differs between situations, exponentiated by exponent *t*, which is less than 1. Each additional source (*N*) increases the degree of impact, however to a lower degree than the (*N*-1)th source. Thus, the fourth actor has less impact than the second one. The presence of the exponent less than one points to a concave relationship between majority size and the degree of social impact or social influence, pointing to a square root function.

This hypothesis is strengthened by another model showing a concave relationship between majority size and the probability of following the majority (MacCoun, 2012), which is the social influence model (Tanford & Penrod, 1984). This model was based on a computer simulation of jury decisions that uses different group sizes in order to differ between the number of social sources and their targets. The model incorporates both majority and minority processes and shows a nonlinear relationship between the size of majority groups and their social influence on group members. This model proved fitting in a meta-analysis on several social influence studies, mainly based on the Asch paradigm (Bond, 2005). The level of social influence obtained is based on how many targets and sources are present in a certain social situation. When the number of targets is a minority of 1, this results in the following equation, based on a revised version of the social influence model (Coultas, 2004; MacCoun, 2012):

$$I = \exp(-4 \times \exp(-N^{1.075}))$$

I = impact and *N* = number of sources of influence

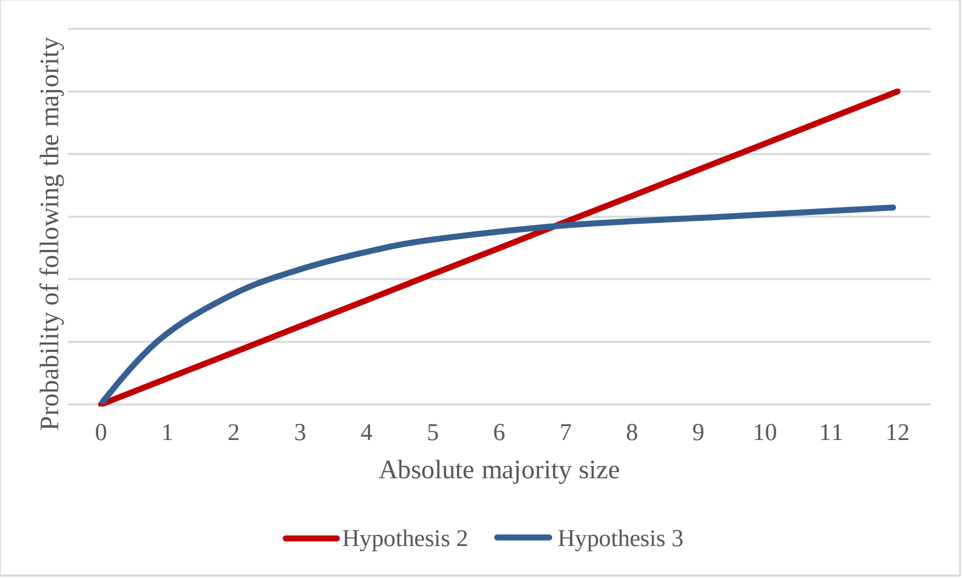
This results in the situation where the second and third decision-maker have more social influence than the first, but this increase in social influence becomes less strong at a certain point. This leads to an S-shaped function. Based on these theories the following hypothesis can be derived:

Hypothesis 3: Majority size has a concave impact on the probability of following the majority

2.3 Graphical display of the predicted mathematical forms of the relationship between majority size and the probability of following the majority

The predicted mathematical forms of the relationships between absolute majority size and the probability of following the majority might be slightly unclear without a visual representation. Therefore, figure 1 shows a broad estimate of what these mathematical forms look like based on the before mentioned hypotheses. The red line shows the predicted effect of hypothesis 2, predicting a linear effect of absolute majority size on the odds of success of following the majority. The blue line shows the predicted effect of hypothesis 3, predicting a concave effect of absolute majority size on the odds of success of following the majority.

Figure 1: the mathematical forms of the relationships between majority group size and the probability of following the majority based on hypothesis 2 and 3



3. METHODS

3.1 Experimental design

<u>Procedures</u>

The data used in this research has been collected in a laboratory experiment conducted by prof. dr. Arnout van de Rijt and prof. dr. Vincenz Frey at Utrecht University. In this experiment groups ranging from 10 to 14 participants were faced with 30 binary questions. The participants answered the questions in an anonymous computer set up, where they were unable to see each other and had no access to the Internet to find the correct answers to the questions. These questions were placed into five different categories, each containing six questions. The categories were visual, art, equations, history and geometry. Participants had 20 seconds for every question, in which they had to choose between alternative A and B, of which one was the correct answer. A count down time was shown on their screens to be aware of the remaining time. Participants could also refuse to answer by letting the 20 seconds pass without answering the question.

The experiment was set up so that subjects had to answer the questions sequentially, without having to wait for others in the group. Every subject answered the questions in the same cyclic order, but was appointed to a different starting position in this cycle. This way every subject had to answer a question at the same time. The experiment consisted of two conditions, the social influence condition and the independent condition. Subjects were randomly appointed to one of the conditions. In the social influence condition, participants were truthfully shown how popular both options were, by showing how many others prior to them had chosen option A and how many had chosen option B. In the independent condition however, these popularity counts were not shown to the participants.

The experiment used an incentive scheme, based on common organizational decision-making settings in which the team outcome is most important, but individual choices are also taken into account. Therefore, every subject got €0.10 for each question the majority answered correctly and another €0.05 for each individual correct answer. This design focused on majority outcomes and was implemented to strengthen social influence by showing subjects the potential harmful effects of individual choices on group accuracy.

Subjects

The experiment consisted of eight experimental sessions, involving 192 participants in total. Each of the 192 participants answered 30 questions, leading to a final *N* of 5670 choices. In each session the total number of participants, 21, 24 or 27, were randomly assigned to two groups of comparable size. These groups were randomly assigned to either the control or the social influence condition. The distribution of the amount of subjects in the control condition and social influence condition varied between sessions, because of differentiation in the amount of total subjects per session, which led to group sizes from 10 to 14. On average, slightly more people were placed in the control condition. When performing the analysis the social influence condition was used, because only this condition allowed for the interpretation of the predicted effects. The independent condition was used as control condition to compare results and make sure no incorrect social influence effect seemed to appear.

The sample consisted mostly of undergraduate students from 18 to 30 years old, with an average of 24. There were just a couple of respondents between the age of 30 and 60. The subjects came from many different educational backgrounds, from social sciences to medicine. The majority however, came from a social sciences or economic background. A small majority of the subjects were born in the Netherlands, but many came from other countries as well. The majority of subjects were female (67%), leading to an underrepresentation of men. However, the percentage of women enrolled at Utrecht University is higher than the percentage of men, especially in the social sciences (CBS, 2018). This combination makes the sample representative for Utrecht in general.

3.2 Operationalization of the variables

Dependent variable

The dependent variable in this research is *following the majority*. As described in the theory section, people follow the majority, indicating that they are influenced, when they purposely choose the same answer as the majority before them did. Therefore, the variable measuring conformity shows if the answer of individual subject is the same as the answer given by the majority. This was done by creating a sensor variable that shows if the majority either chooses option A or option B. When there was no majority at that choice opportunity, the sensor variable randomly chooses A or B with a 50 percent likelihood[2]. This was done to maintain as much data as possible and only had a minor, non-significant impact on the results. Another variable was created that showed if this sensor variable was equal to the answer chosen by the majority group or not. This resulted in a discrete binary variable, in which a score of '0' showed that the answer of the subject was different from the majority group and a score of '1' showed that the subject chose the same answer as the majority group. The cases where no answer was given within the 20 seconds were filtered out, because they included no useful information for the analysis.

Independent variables

The main independent variable in this research is majority size. As shown in the literature section, the urge to follow the majority is expected to depend on its size. However, the hypothesis show different relationships between majority size and following the majority. In order to test these contradicting hypotheses, multiple majority size variables were created.

The basic variable *majority size* was created to test hypothesis 2. This variable measured the number of individuals in the majority group at every choice opportunity. The variable was created by subtracting the number of times option A was chosen, from the number of times option B was chosen by all subjects. This variable was transferred into an absolute variable to allow for easy interpretation. This results in a scale from 0 to 12, where 0

[2] This randomized variable led to minor differences in the results. However, this did not have significant impact and can therefore be accepted.

represents no majority and 12 represents a majority size of 12 subjects. The further away from 0 the more popular the answer was at that time.

In order to test hypothesis 3, predicting a concave relationship between majority size and the degree of social influence, an additional variable on majority size *majority size square root* was created. The third hypothesis predicted that when majority size increases the probability of following the majority increases, but less for every additional person in the majority group. Therefore, a variable was needed that was able to show this predicted function. The square root of absolute majority size allowed to interpret this relationship, because it has the properties of being strictly increasing and concave (Kennan, 2001). The variable was created by using the square root term of the number of times option B was chosen minus the number of times option A was chosen. This resulted in a scale from 0 to 3.46, where 0 predicted the square root of a majority of 0 and 3.46 predicted the square root of a majority of 12.

Furthermore, in order to test hypothesis 1 and to allow for an explorative test of the relationship between absolute majority size and the probability of following the majority, dummy variables were created for every majority size. There was a dummy for every absolute majority size ranging from 0 to 12, leading to the creation of dummies '*m0'* to '*m12'*.

<u>Control variables</u>

The control variable in this research was question difficulty. Since the questions ranged in difficulty it was necessary to check for these various levels of difficulty. When answering a less difficult question, subjects could have been expected to follow the correct majority based on their knowledge of this answer being correct, instead of following the majority based on social influence. This could lead to an misinterpretation of the effect of majority size on following the majority, both in the social influence condition and the control condition. Furthermore, Lucas et al. (2006) showed that the probability of following the majority can increase when faced with more difficult tasks due to lower self-efficacy. Therefore, it was necessary to control for the effects of every individual question to control for the differing difficulty. This was done by creating dummies for every question that controlled for the tension of the majority to choose either A or B at every choice opportunity, based on the sensor variable. This type of controlling worked through a fixed effects method. In this fixed effects model the mean for every question is a group-specific fixed quantity, meaning that the mean for every question difficulty is set. This resulted in dummies *qn1* to *qn30*, each representing a dummy for one of the 30 questions.

3.3 Methods

The contradicting nature of the before mentioned hypotheses allows for several tests. First of all an explorative model was created which showed the odds of following the majority for every majority size. These results were shown in a graph of the log odds to allow for easier interpretation, which will be explained in detail later on. To be able to derive which predicted relationship had the best fit for our data, two models were used. These models were compared based on evaluation criteria to validate which hypothesis could be confirmed and which could be rejected. When performing the analysis the social influence condition was used, because only this condition allowed for the interpretation of the predicted effects. The independent condition was used as control condition to validate that no incorrect social influence effect appeared.

Because this article focusses on the mathematical form of the relationship between majority size and the probability of following the majority, an explorative logistic probability model was used first to give initial insight in which hypothesis could be rejected and which needed further analytical inspection. This model also allowed for testing hypothesis 1. By using this model, insight in the potential shape of the relationship between majority size and social influence was provided. This model used dummies *m1* to *m12*, for every majority size between 1 and 12 as independent variables, with a baseline dummy *m0*, representing a majority of 0. The dependent variable measuring the degree of social influence was *following the majority*. Where 1 was equal to following the majority and 0 was equal to not following the majority. The control variables used in this model were dummies *qn1* to *qn30*, representing a dummy for every question. Using the fixed effects of every question, allowed to control for factors influencing the results such as individual question difficulty. Since all variables were of binary nature, the proper method of analysis was a binary logistic regression. This analysis gives the estimated likelihood of belonging to one of the binary groups of the dependent variable in terms of odds ratios (Hilbe, 2009). In the theory section probability was used to describe the effect of majority size on following the majority, which is easy to retrieve from a binary logistic regression. This made using this type of regression even more suitable for the interpretation of this relationship. Based on these results a graph representing the logarithm of odds of following the majority for every majority size between 1 and 12 was created, allowing for an easier interpretation of the predicted mathematical forms. The model from which the graph was retrieved was used to test hypothesis 1.

This explorative model was followed by two logistic probability models to test hypothesis two and three, which were shown in one table. These models used the same dependent variable and control variables, but the independent variables were changed to variables of categorical nature as mentioned below. This combination of variables still required a binary logistic regression (Hilbe, 2009). Therefore, a binary logistic regression model is used for testing all hypotheses. The first of these models tested the existence of a relationship between majority size and conformity, thereby testing hypothesis 2. While the dependent variable and control variables remained the same as in model 1, the independent variable was changed to *majority size*. The second model, allowed to test for a concave relationship between majority size and conformity, thereby testing hypothesis 3. The independent variable in this model was changed to *majority size square root*, while maintaining the same dependent and control variables as in the other models.

When using several models testing contradicting hypotheses, it would have been incorrect to use a nested model. However, it was still necessary to decide which model had the best fit. In order to select the model with the best fit for the relationship between majority size and conformity, both Bayesian information criterion and Akaike information criterion were used. These model selection criteria are good ways of comparing the fit of non-nested models (Posada & Buckley, 2004). The criteria indicate how much structural information is modeled in the data, such as the structure of relationships and variance, and separate this information from noise (Burnhem & Anderson, 2004). The model with the lowest score is able to provide the most structural information of the data and can therefore be chosen as the best fit.

3.4 Descriptive statistics

The descriptive statistics of the variables that have been used in the analysis are shown in table 1. The average score on the dependent variable, *following the majority,* is .78 in the social influence condition. As expected, the score on this dependent variable in the control condition was lower with an average of .62. This variable ranged between 0 and 1, making this a binary variable. The average of the independent variable *majority size* was 3.66 in the social influence condition and 2.59 in the control condition. The minimum of this variable was 0 and the maximum was 12. The average of the other independent variable *majority size square root* was 1.68 in the social influence condition and 1.37 in the control condition. This variable ranged from 0 to 3.46. The dummies representing the fixed effects for every question ranged from -1 to 1. The median scores and standard deviations are not given in this table, because they differ for every question.

Table 1 Descriptive statistics

	Social influence condition			*Control condition*			
Variables	*N*	*Mean*	*S.D.*	*N*	*Mean*	*S.D.*	Range
Following the majority	2681	.78	.41	2791	.62	.49	0 - 1
Majority size	2820	3.66	2.93	2940	2.59	2.37	0 - 12
Majority size square root	2820	1.68	.92	2940	1.37	.84	0 − 3.46
Question difficulty dummies qn1 to qn30	2820			2940			-1 − 1
Valid N	2681			2791			

Note: *M* and *SD* are not given for the question difficulty dummies, due to different scores for every dummy.

4. RESULTS

4.1 Results of the logit function of the effects of every majority size between 1 and 12 on the odds of following the majority

In order to give initial insight in the relationship between majority size and the odds of following the majority a binary logistic regression analysis was performed. This regression showed the effects of every majority size between 1 and 12 on the probability of following the majority compared to a majority size of 0, while controlling for the fixed effects of every question. The results of this analysis are shown in appendix A. This model provided a good fit for the data, since it was statistically significant in comparison to the null model $\chi^2(42) = 464.13$, $p < .001$. The model was able to explain 24.5% of the variance in following the majority and was able to classify 81.2% of the cases correctly. This good fit indicated that the model could be used for initial interpretation of the predicted results.

The table of this model was not presented in the results section, because it was more intuitive to present the results in a graph showing the logarithm of the odds of success and their 95% CI interval for every majority size from 1 to 12. Unfortunately, there was not enough power in the last categories, majority sizes 8 to 12, to accurately fit the mathematical form of the predicted relationship. Therefore, the graph in figure 2 shows the logarithm of the odds of following the majority and the log odds of the 95% CI intervals for every majority size from 1 to 7. However, the graph showing the logarithm of odds for every majority size from 1 to 12 was shown in appendix B for completeness. The graph allowed to see the mathematical form of the odds more clearly and made it possible to plot the predicted linear and concave functions in it. The logarithm of the odds of following the majority were used, which created a symmetry around zero, allowing for easier interpretation than regular odds (Jaccard, 2001). Based on this graph it was possible to make initial predictions about the hypotheses regarding relationship between majority size and the probability of following the majority.

To make sure that these effects were due to social influence and no mistakes were made in the fixed effects method, the same analysis were performed on the control condition. These results are shown in appendix C and indicated that no significant effects of majority sizes were found in the control condition. The log odds of following the majority for every majority size were nearly 0, with corresponding odds around 1, indicating that there was a 50% chance of following the majority. This showed that there was no inclination towards following the majority in the control condition. Therefore, we can conclude that the main model accurately predicts the degree of social influence and can be used for the validation of the hypotheses.

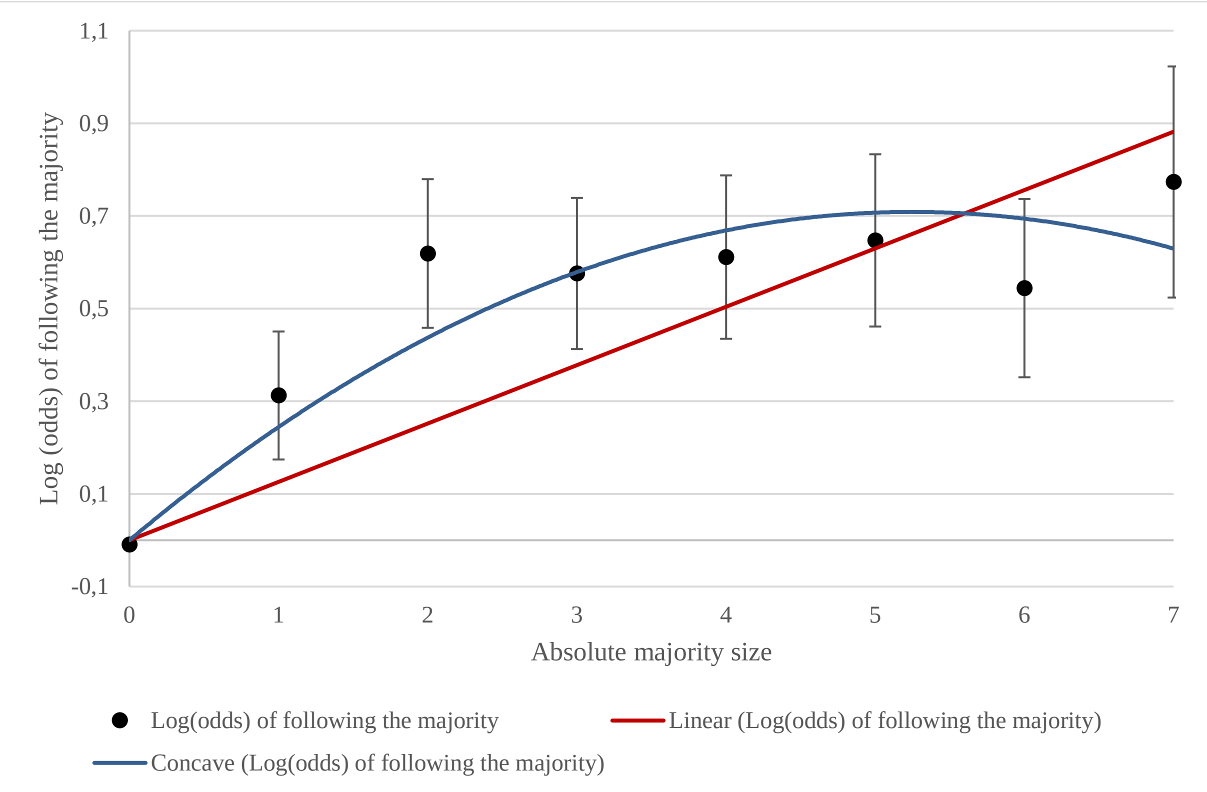
The first hypothesis predicted that the probability of following the majority group depends on the absolute size of the majority group. If this hypothesis was correct, there would be no effect of following the majority when there is no majority, indicating a 50% chance of following the majority. There would have also been a significant effect on following the majority, when the majority size increased from 0 to 12. As shown in figure 2 and appendix A, the constant is negative and non-significant (b = -.021, p = .852, OR = .979). This indicates that when there is no majority the odds of following the majority was around 1, showing no effect on following the majority. As shown in appendix A, for every incremental step in majority size up till 11 the probability of following the majority increased significantly. Based on these results, hypothesis 1 can be confirmed.

The second hypothesis predicted that majority size has a linear impact on the probability of following the majority. When looking at the fitted linear function in figure 2, it becomes clear that the log odds of the majority sizes 1 to 4 were above the linear function, predicting a higher probability of following the majority. While a majority size of 5 seemed to be in line with the predicted function, the remaining majority sizes of 6 and 7 had log(odds) lower than the predicted log(odds) of following the majority. When looking at the fitted linear function in appendix B, representing the model with majority sizes from 1 to 12, the linear fit did not seem to correlate with the data either. As majority size increased incrementally, the log odds of following the majority did not seem to increase exponentially. Based on this distribution there was little initial support for hypothesis 2.

The third hypothesis predicted that majority size has a concave impact on the probability of following the majority. When looking at the fitted concave function in figure 2, it becomes clear that the log odds of most majority sizes and their corresponding log odds 95% confidence intervals fitted the function. The fit of the concave function in appendix B seemed to correspond with this outcome and showed that the concave line fell within most of the 95% confidence intervals for the majority sizes of 1 to 12. The probability of following the majority seemed to increase rapidly at first, but slowed down for majority sizes bigger than 2. Based on this distribution there seemed to be initial support for hypothesis 3.

Concluding, the concave function of majority size seemed to have a better fit for the probability of following the majority. However, in order to validate the proposed hypotheses it was necessary to officially analyze the predicted mathematical functions and compare the corresponding models. The results of this analysis are shown in the next section.

Figure 2: the log odds of following the majority for every majority size between 0 and 7 and their 95% CI



4.2 Results of the binary logistic regression models predicting the linear and concave relationships of majority size and the probability of following the majority

In order to officially test hypothesis 2 and 3, two binary logistic regressions were performed, showing the linear and concave functions and their corresponding effects of majority size on the probability of following the majority. The results of these binary logistic regressions are shown in table 2. Initially, both models seemed to provide a good fit. The linear model was statistically significant in comparison to the null model $\chi^2(31) = 404.536$, p <.001. The model was able to explain 21.6% of the variance in following the majority and was able to classify 80.9% of the cases correctly. The concave model was also statistically significant in comparison to the null model $\chi^2(31) = 437.598$, p <.001. This model was able to explain 23.2% of the variance in following the majority and was able to classify 81.3% of the cases correctly. Based on these results, both models allowed a good fit for the data and permitted analysis.

The results of the analysis for hypothesis 2 are shown in table 2. There was a positive significant effect of increasing majority size on the probability of following the majority when controlling for the fixed effects of every question, $b = .181$, $p <.001$, $OR = 1.198$ (95%CI: 1.149, 1.250). As the majority size increased incrementally by one subject, the odds of following the majority increased by 1.198 times. This would mean that an increase in majority size by 1 meant that the probability of following the majority increased by $\frac{1.198}{1+1.198} = .545$. The more the majority size increases, the more people are likely to follow the majority, showing a linear impact of majority size on the probability of following the majority. Based on these results, there seemed to be support for hypothesis 2. However, this predicted relationship still had to be compared with hypothesis 3, based on information criteria.

The results of the analysis for hypothesis 3 are shown in table 2. There was a positive significant effect of increasing the square root of majority size on the probability of following the majority when controlling for the fixed effects of every question (b = .612, p <.001, *OR* = 1.845 (95%CI: 1.644, 2.070). As the majority size increases incrementally by one subject, the odds of following the majority increase by 1.845 times. This would mean that an increase in majority size by 1, $(\sqrt{1} = 1)$, means that the probability of following the majority increases by $\frac{1.845}{1+1.845} = .648$. This indicates that the inclination to follow the majority first increases rapidly when the majority size becomes bigger and then starts to slow down, showing a concave impact of majority size on the probability of following the majority. Based on these results, there also seemed to be support for hypothesis 3.

Table 2: the results of the binary logistic regression of the effect of linear and concave impact of absolute majority size on following the majority

	Linear		*Concave*	
Variable	*B*	SE*B*	*B*	SE*B*
Majority size	.181***	.021		
Majority size square root			.612***	.059
Constant	.476***	.077	.122***	.096
Df	31		31	
-2 Log Likelihood	2398.249		2365.187	
BIC	2642.96		2609.90	
AIC	2460.25		2427.97	
N cases	2681		2681	

Note1: The dependent variable in this analysis is social influence coded so that 0 = did not follow the majority answer and and 1 = followed the majority answer. Note2: Controls are dummies qn1 to qn30 (omitted from the table).

Both effects showed to be positively significant, which made it was necessary to compare the models on both Akaike information criteria (AIC) and Bayesian information criteria (BIC). These criteria make it possible to compare non-nested models with equal sample sizes and are based on the log likelihood of the model and parameters of the model. The lowest score indicates the best fit, which shows the most structural information for the data. The difference between the two gives an indication of how much better the model is. The AIC is less strict than the BIC, because it does not take into account the number of observations. The scores on both criteria for the two models are shown in table 2. The scores for the AIC were calculated from the following equation (Statistica, 2013):

$$\text{AIC} = -2LL + 2p \text{ (number of parameters)}$$

$-2LL$ = log likelihood, p = number of parameters in model

As shown in table 2, the AIC of the concave function (AIC = 2427.97) resulted into a lower score than the AIC of the linear function (AIC = 2460.25). This indicated that the concave function was able to predict more of the structural information in the data and had a better fit. The difference between the two AIC scores was very high, with a score of 32.28. This indicates that the concave function had a much better fit on the Akaike information criteria.

The scores on the BIC were calculated from the following equation (Statistica, 2013):

$$\text{BIC} = -2LL + p\ (number\ of\ parameters) * ln(N)$$

-2LL = log likelihood, *p* = number of parameters in model, *N* = number of observations

As shown in table 2, the concave function still seemed to have a better fit when looking at the BIC score (BIC = 2609.90) in comparison to the linear function (BIC = 2642.96). When controlling for sample size, the concave function was still able to predict more of the structural information of the data with a BIC difference of 33.06.

Based on the information criteria, it became clear that the concave function had a much better fit for the data than the linear function. This finding was supported by the logit function showing the log odds of following the majority for every majority size to 7 in figure 2. With an additional positive significant effect of the square root of majority size on the probability of following the majority, hypothesis 3 was confirmed. Based on these results it became clear that hypothesis 2 could be rejected.

5. CONCLUSION AND DISCUSSION

5.1 Conclusion

This study aimed at providing more insight in the long debated relationship between majority size and the probability of following the majority, focusing on the mathematical form of this relationship. To provide more clarity on this matter experimental data was used that showed the tendency to follow the majority in sequential binary decision-making when provided with information on the choices of a person's predecessors. In order to find out if a relationship between majority group size and the probability indeed existed, the effects of majority group size on following the majority were tested first. To provide more insight on the mathematical form of this relationship, tests on the linear impact and the concave impact of majority size on the probability of following the majority were performed.

Hypothesis 1 predicted that the probability of following the majority group depended on the size of the majority group. This hypothesis was based on several major theories explaining this relationship based on informational influence and/or normative influence (Asch, 1955; Latané, 1981; Tandord & Penrod, 1984). Others showed that this was due to positive feedback of initial inequalities (DiPrete & Eirich, 2006; Barabasi & Albert, 1999) or rational choice behavior (Bikhchandani et al., 1998). This predicted relationship between majority group size and the probability of following the majority showed to be significant. This indicates that once someone is in a binary decision-making setting, the probability of following the answer of the majority group and ignoring one's personal information depends on the size of this majority group. If an option changes in popularity, so does the probability of ignoring one's personal beliefs and following the majority. However, when accepting this hypothesis it was not possible to predict which theory was best to explain these results. The social influence that occurred when following the majority could have been due to both normative and informational influence due to the incentive scheme that was used, which increased the group focus. However, this experiment did not allow for clearly testing these different mechanisms.

Hypothesis 2 predicted that majority size has a linear impact on the probability of following the majority. This would mean that for every additional person in the majority group the probability of following this majority in their perceived opinion or belief increased exponentially. The cumulative advantage model explained that this exponential increase in following the majority answer, is due to ongoing positive feedback that comes from popularity in early stages (DiPrete & Eirich, 2006; Huber, 1998). When the majority answer becomes more popular, the probability of choosing this answer increases. Barabási and Albert (1999), showed that a preferential attachment mechanisms exists in networks, which means that the more popular an option is, the more likely it is to increase in popularity. Newcomers perceive the popularity of a choice and are more likely to follow it. However, this linear impact of majority size on the probability of following the majority was not found to be significant, leading to the rejection of hypothesis 2. The fact that hypothesis 2 was rejected, indicated that the cumulative advantage model and the preferential attachment mechanism did not apply to this type of decision-making. This could be explained by the experimental design, which was more focused on short-term group dynamics than the positive feedback an answer or person receives on the long term, which has been the focus of most research on these mechanisms. Possibly, exponential growth in this type of social influence only occurs on the long term. Furthermore, these models mostly occur in scale-free networks, while the network used in this experiment did have clear boundaries.

Hypothesis 3 predicted that majority size has a concave impact on the probability of following the majority. This indicates that the first people creating a majority group increase the probability of following them in their beliefs or opinions more than the people that will join the majority group later on. Latané (1981) described this concave impact based on social impact theory. This theory explains that the initial majority of sources of social impact strengthens the belief that the chosen option is the right one and that they have more power to punish norm deviation more than the additional ones. Tanford & Penrod (1984) explained this concave relationship based on their social influence model, which shows that both normative influence and informational influence increase when the majority becomes bigger, but less for every additional person. In this research this concave impact of majority size on the probability of following the majority was found to be positively significant. Therefore hypothesis 3 was confirmed, indicating that the social influence of the majority group increases rapidly when the majority increases in size at first, but increases much slower for additional members of the majority group. This led to support for the proposed mechanisms of the social impact theory and the social influence model. Even though, when looking at figure 2 and appendix C, this would point more towards the R-shaped function as predicted by social influence theory than the S-shaped function predicted by the social influence model, this research does not have the capacity to indicate which is most fitting.

Based on these findings, it can be concluded that majority group size is an important factor when looking at social influence based on the probability of following the majority. It was found that this probability differs between majority sizes and is stronger for the initial members of the majority group than those that join the majority group later on. Therefore, the debate between the mathematical form of the effect of majority size on the probability of following the majority seems to be answered by the better fit of the concave function. This shows that the social impact model as well as the social influence theory seem to be fitting in anonymous decision-making. This gives interesting insights in the decision-making process in other cases of choosing between alternatives in anonymous decision situations, such as voting-polls or choosing between two products.

As mentioned in the introduction, the spread of erroneous information due to undermining mechanisms of the wisdom of the crowd, is something that happens regularly. The findings of this paper can be helpful when trying to decrease the spread of erroneous information on for example social network sites, in juries or organizational settings. First of all, popularity rates could not be shown at all, because when there is no clear majority there is no inclination towards one option based on social influence. Another option is trying to intervene with the initial sources of incorrect information since these first sources of the majority, providing the incorrect answer, have the most impact on others who are making a decision. The findings can also be helpful when trying to improve health or consumption behavior by using social norms. If people perceive that a couple of others have chosen the option that you want them to choose, this will often be sufficient to let them make the same choice. It does not have that much more impact to use very big majority sizes, because they are less important to deviating from someone's personal beliefs. These findings supported other studies indicating that merely showing that the majority of people have chosen a certain option is sufficient to induce the desired behavior (Goldstein, Cialdini & Griskevicius, 2008).

5.2 Discussion

Even though this research led to some interesting insights in the relationship between majority size and the probability of following the majority, just as any other scientific paper it also had its flaws. In this section some of these flaws will be elaborated on and the potential for future research based on the results will be given.

Originally, one of the benefits of this experimental design was that it permitted to look at bigger majority sizes than those used in previous research. The data allowed to look at majority sizes up to 12. However, in a setting with a maximum of 14 subjects these bigger majority sizes of course did not occur that often, but when they occurred, they did not have impact on that many other subjects. This resulted in a lack of power when looking at the fit of the mathematical relationship of majority size on the probability of following the majority for these bigger majority sizes. Therefore, the main explorative graph does not show the log odds of following the majority for majority sizes bigger than 7. This might have influenced the shape of the relationship slightly and having more power might have led to a better fit of the mathematical function on the data. Therefore, future research focusing on the importance of majority size on social influence should incorporate bigger sample sizes.

It was shown that the absolute majority size effects the probability of following the majority, but relative majority size might have been even more relevant for explaining this relationship. Especially, since the graph represented in figure 2 seems to have some overlap with the predicted shape of the Other Total Ratio (Mullen, 1987) as presented by Bond (2005), which predicts that relative size of the majority group is a better indicator of following the majority. However, due to lacking clarity about the effect of absolute majority size on the probability of following the majority, this research was a good first step. Furthermore, additional research on this matter is currently being done by Veldkamp (2019), which might lead to an interesting addition to the results of this paper.

Even though, the concave function showed to be significant, the level of analysis in this paper did not allow for testing which of the theories predicting this relation had the best fit. Additional analysis about the fit of both the social influence model and the social impact theory in anonymous decision-making might be necessary to unravel which model has the best fit. This flaw relates to a more general flaw in this research, which entails that due to the mathematical approach the micro-macro mechanisms behind following the majority based on its size have not clearly been found. However, most research on this topic tends to have a focus towards a mathematical approach instead of trying to explain the deeper meaning of the process of following the majority based on its size. Furthermore, the data set did not allow for testing potential micro level explanations, such as authority or group-identity, which have been shown to have effect on this process in conformity experiments (Schöbel, Rieskamp & Huber, 2016;). This indicates the demand of future research on the complicated topic of following the majority based on majority size.

Therefore, an interesting option for future experimental research would be creating an online anonymous binary-choice experiment which uses bigger sample sizes to give a better understanding of the effects of bigger majority sizes. This experiment should also include a simulation of group identities or authority among subjects in order to test their impact on the probability of following the majority based on its majority size. Creating more insight in these matters in anonymous decision-making could be beneficial for understanding the micro-macro links of the relationship between majority size and the probability of following the majority.

Furthermore, both Asch's (1955) findings and the literature on informational cascades (Bikhchandani et al., 1998), predict a threshold function of the effect of absolute majority size on the probability of following the majority. Due to lacking power it was not possible to correctly test this function. However, when looking at the figure 2 there does seem to be a change in the probability of following the majority around a majority size of 2. Diving further into the thresholds of this relationship might be interesting for future research.

As shown in this research, every decision influences the outcomes in life. Even though, some of the parameters could have been optimized the results have left us with some interesting insights in the complex relationship between majority size and the probability of following the majority. The size of the absolute majority size has effect on the probability of following the majority and this probability follows a concave function. Hopefully, the proposed ideas for future research will lead to more people making the choice to unravel this phenomenon of social influence.

6. REFERENCES

Asch, S.E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men*(pp. 177–190). Pittsburgh, PA: Carnegie Press.

Anderson, L. R., & Holt, C. A. (1996). Classroom games: Information cascades. *Journal of Economic Perspectives, 10*(4), 187-193.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives, 12*(3), 151-170.

Bond, R. (2005). Group size and conformity. *Group processes & intergroup relations*, *8*(4), 331-354.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research, 33*(2), 261-304.

Centraal Bureau voor Statistiek. (2018). *Jonge-vrouwenoverschot in veel universiteitssteden.* Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2018/34/jonge-vrouwenoverschot-in-veel-universiteitssteden

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55, 591-621.

Crutchfield, R. S. (1955). Conformity and character. *American Psychologist, 10*(5), 191

Cukierman, A. (1991). Asymmetric information and the electoral momentum of public opinion polls. *Public Choice*, *70*(2), 181-213

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annu. Rev. Sociol., 32*, 271-297.

Fischer, R., & Vauclair, C. M. (2011). Social influence and power. A Psicologia Social Brasileira: Principais temas e vertentes, *January2015*, 1-35.

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature, 75*(7), 450-451.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research*, *35*(3), 472-482.

Hilbe, J. M. (2009). *Logistic regression models*. Chapman and hall/CRC.

Huber, J. C. (1998). Cumulative advantage and success-breeds-success: The value of time pattern analysis. *Journal of the American Society for Information Science, 49*(5), 471-476.

Jaccard, J. (2001). *Interaction Effects in Logistic Regression, Issue 135.* SAGE.

Kennan, J. (2001). Uniqueness of positive fixed points for increasing concave functions on Rn: An elementary result. *Review of Economic Dynamics*, 4(4), 893-899.

Latané, B. (1981). Psychology of social impact. *American Psychologist, 36*, 343–356.

Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review, 88,* 438–45

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F. & Schudson, M. (2018). The science of fake news. *Science, 359*(6380), 1094-1096.

Lopez-Pintado, D., & Watts, D. J. (2008). Social influence, binary decisions and collective dynamics. *Rationality and Society, 20(*4), 399-443.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the national academy of sciences, 108*(22), 9020-9025.

MacCoun, R. J. (2012). The burden of social proof: Shared thresholds and social influence. *Psychological review*, *119*(2), 345.

Marteau, T. M., Ogilvie, D., Roland, M., Suhrcke, M., & Kelly, M. P. (2011). Judging nudging: can nudging improve population health?. *Bmj*, *342*, d228.

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *isis*, *79*(4), 606-623.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review*, *26*(1), 67-82.

Mullen, B. (1983). Operationalizing the effect of the group on the individual: A self-attention perspective. *Journal of Experimental Social Psychology*, *19*, 295–322.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological review*, *97*(3), 362.

Oxford Dictionary of Psychology (3th ed.). (2014). Oxford: Oxford University Press.

Perc, M. (2014). The Matthew effect in empirical data. *Journal of The Royal Society Interface*, *11*(98), 20140378.

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology*, *53*(5), 793-808.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *science*, *311*(5762), 854-856.

Schöbel, M., Rieskamp, J., & Huber, R. (2016). Social influences in sequential decision making. *PloS one*, *11*(1), e0146536.

Statistica. (2013). *STATISTICA Formula Guide: Logistic Regression.*

Tanford, S., & Penrod, S. (1984). Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin, 95*(2), 189.

Van de Rijt, A., Kang, S. M., Restivo, M., & Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 201316836.

Veldkamp, J. (2019). Social influence: A comparison between the absolute and relative component of majority size in their effects on conformity (Unpublished bachelor thesis). Utrecht University, The Netherlands.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *science*, *330*(6004), 686-688.

Young, H. P. (2009). Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American economic review, 99*(5), 1899-1924.

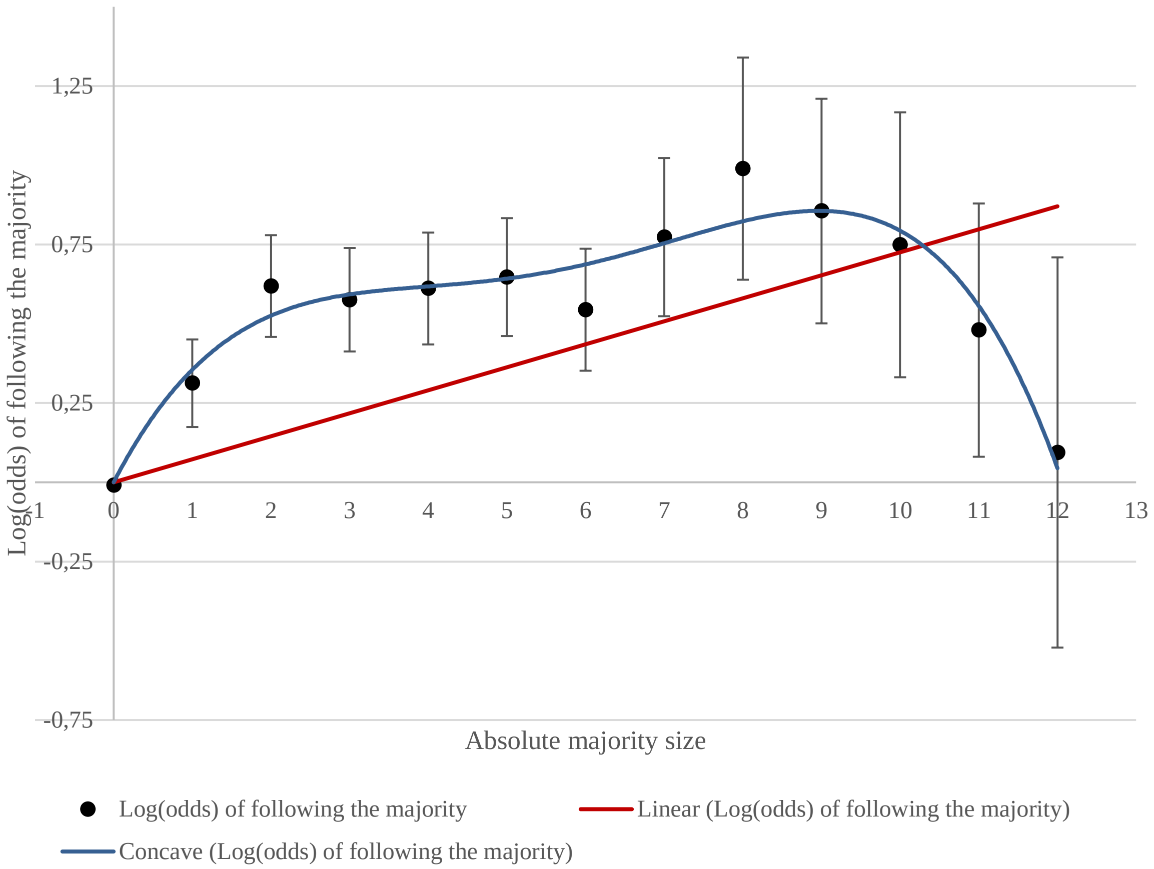
7. APPENDICES

Appendix A: the results of the model of the probability of following the majority for every dummy of majority size in the social influence condition

Predictor	*B(SE)*	OR *e*B	95% CI for OR Lower	Upper	*P*
Ms1	.720(.162)	2.053	1.491	2.823	<.001
Ms2	1.425(.189)	4.160	2.874	6.020	<.001
Ms3	1.326(.192)	3.766	2.586	5.483	<.001
Ms4	1.408(.207)	4.086	2.722	6.134	<.001
Ms5	1.491(.218)	4.439	2.893	6.821	<.001
Ms6	1.253(.226)	3.502	2.248	5.454	<.001
Ms7	1.781(.293)	5.934	3.341	10.541	<.001
Ms8	2.278(.412)	9.759	4.355	21.869	<.001
Ms9	1.970(.416)	7.172	3.173	16.211	<.001
Ms10	1.725(.491)	5.614	2.145	14.694	<.001
Ms11	1.105(.469)	3.020	1.204	7.578	.019
Ms12	.217(.723)	1.242	.301	5.124	.765
Constant	-.021(.115)	.979			.852
-2 Log Likelihood	2338.66				
N cases	2681				

Note1: The dependent variable in this analysis is social influence coded so that 0 = did not follow the majority answer and and 1 = followed the majority answer. Note2: Controls are dummies qn1 to qn30 (omitted from the table). eB = exponentiated B.

Appendix B: the results of log(odds) function of the effects of every majority size between 1 and 12 on the log(odds) of following the majority



Appendix C: the results of the model of the probability of following the majority for every dummy of majority size in the control condition

Predictor	*B(SE)*	OR *e*B	95% CI for OR Lower	Upper	*P*
Ms1	.095(.131)	1.099	.851	1.420	.468
Ms2	.068(.144)	1.070	.807	1.419	.637
Ms3	-.096(.160)	.908	.664	1.242	.546
Ms4	.108(.188)	1.114	.770	1.611	.567
Ms5	-.028(.206)	.972	.649	1.455	.891
Ms6	-.287(.230)	.750	.478	1.177	.211
Ms7	-2.227(.271)	.797	.469	1.356	.403
Ms8	-2.222(.352)	.801	.402	1.597	.529
Ms9	.234(.469)	1.263	.504	3.168	.618
Ms10	-.608(.507)	.544	.202	1.470	.230
Ms11	-.385(822)	.680	.136	3.409	.640
Constant	-.085(.098)	.918			.385
-2 Log Likelihood	3244.409				
N cases	2791				

Note1: The dependent variable in this analysis is social influence coded so that 0 = did not follow the majority answer and and 1 = followed the majority answer. Note2: Controls are dummies qn1 to qn30 (omitted from the table). Note3: the absolute majority size of 12 is not shown in the table, because it did not occur in the control condition. eB = exponentiated B.