May
2018

Crowd Predictions in Venture Capital

Authors:
Oscar Kehlet
Cand. Merc. International Business
60751

Anders Scheuer Rasmussen
Cand. Merc. Strategic Market Creation
82186

Supervisor
Carina Antonia Hallin

CBS COPENHAGEN BUSINESS SCHOOL HANDELSHØJSKOLEN

Number of Characters: 189.349 / 224.578
Number of Pages: 106

Abstract

The objective of this study was to apply *The Wisdom of Crowds* in the setting of Venture Capital. It was explored whether crowds that were large, diverse and independent were more accurate in their predictions than small groups of experts within the industry of venture capital. Additionally, it was investigated whether certain characteristics of experts could explain predictability and if predictions were a result of confidence bias. To measure to the performance of the crowds, a crowd prediction tournament was developed, where four early-stage start-ups were assessed and evaluated on. The results indicated that a group of 58 non-experts could outperform 7 experts with domain knowledge. However, the results were insignificant. The researchers found significant evidence that, when evaluating the potential of start-ups, the product was the most essential evaluation criteria.

ACKNOWLEDGEMENTS

First of all, we would like to state our appreciation and thank our supervisor who have been extremely helpful during the process. She has been proactive and accessible from the very beginning.
Secondly, our gratitude goes to Almi Invest for providing the historical start-up cases. They played a vital role in the crowd prediction tournament.
Additionally, we would like to thank PreSeed Ventures, Vækstfonden and ByFounders for making it possible to conduct interviews, which provided crucial knowledge in the preface of the study.
Lastly, a huge thanks to all the participants of the crowd tournament who were willing to spend 20 minutes on predicting start-ups to their best ability.

Contents

Bill Gates says the hit rate of venture capital funds is “*pathetic*”. (Nisen, 2014)

Chapter 1 - Introduction

As a result of globalization and digitalization, the rate of new startups is extremely high. In fact, 100 million startups are launched each year (Innmind, 2016). Access to capital, skilful workforce and friendly environments are improving the conditions for entrepreneurs to turn their idea into a global and successful organization. Successful start-ups have large impacts on the economy in terms of creating high-paying jobs, boosting local investment and increasing consumer spending (StartupGenome, 2017). Start-ups in the United States created 2.2 million jobs in 2014 within their first five years of lifespan and in Europe, start-up does on average generate 12.9 jobs within their first 2.5 year in business (Wauters, 2016; More, 2017). Government and politicians constantly strive to improve the conditions for entrepreneurs. Still, most start-ups fail. In fact, nine out of ten start-ups fail (Griffith, 2014). Most start-ups fail due to the fact that there is no need for its product or service in the market (Griffith, 2014). Additionally, if the team fails to utilize their competencies and skills of managing growth within its business, the product or service might be swept away by competitors. Many explanations can lead to failure of a start-up. Despite having a great idea driven by ambitious people, issues such as timing, lack of passion, legal challenges, poor marketing or the threat of running out of cash are all risk factors that can lead to failure of a start-up (Griffith, 2014). A successful start-up is usually determined to be a product of a perfect team and market fit (Appendix 7.2). Yet, in order for startups to grow into successful global companies, they are likely to be in need of capital and advisory. Hence, often most successful start-ups are backed by a venture capital fund.

When venture capitals (VC's) are investing in early-stage start-ups, it is highly related to uncertainty and risk. A venture backed start-up is considered successful when it returns the initial investment 10 times (Lemkin, 2012). Thus, VCs have accepted a business model where only few of the companies within their portfolio are actually going to return the anticipated investment and therefore much capital is "wasted". Never before has so much money fluctuated within VC funds. From 2009 to 2015 the total amount of VC funding in the U.S increased by 430% - from 14,7 billion dollars to 63,3 billion dollars (Meisler, 2016). The same conditions apply for Europe, where the amount of VC invested has increased from 8,6 billion euros in 2014 to 16 billion euros in 2017 (Dealroom, 2017). Even though the industry of venture capital has grown rapidly over the last few years together with the amount of capital available have increased heavily, this has not reflected a significant change in the success rate of venture capital. Based on a study by Harvard Business School senior lecture

Shikhar Ghosh of high potential start-ups in the US, VC's have an average failure rate of 75 % (Gage, 2012), meaning they are not returning the invested capital. If failure is defined as losing the full investment, the failure rate is 30 to 40 %. If failure rates were defined as not returning the projected return of investment, up to 95 % of venture capital backed companies fail (Gage, 2012). An equivalent failure rate would in most other industries be unacceptable; i.e. if a doctor were to lose 30-40 % of his patients, a lawyer only won 25 % of his cases or an intelligence officer were right only in 5 % of his predictions, these would definitely soon be out of work. Nonetheless, in the industry of venture capital this is an accepted business model.

Despite the fact that the rate of start-ups is at its highest and VCs have backed more start-ups due to the higher proportion of available capital in the industry, their success rate has not changed. In other words, VCs have not become smarter or better at predicting which start-ups that eventually will succeed. So why is that? Perhaps it could be a result of not rethinking and developing the business model of venture capital. VC's decision-making process is a perfect example of group work across experts – and group work can indeed be a great tool to solve complex tasks. However, group decisions might also be a great example of biased results that lack diversity (Alper et al., 1998). Thus, group work can undoubtedly also lead to suboptimal decisions.

1.1 Collective intelligence

It is commonly known that two minds think better than one - and ten minds thinks better than two. Pierre Lévy (Lévy, 1997) introduced the term '*collective intelligence*', which states that crowds hold a collective intelligence that foresees future events better than a few experts (Surowiecki, 2004). In 1906, Galton (1907) held a competition. The rules were simple: an ox was going to be weighed and the person who could come up with the closest number to the actual weight of the ox had won. Some of the people among the crowd were experts as they were farmers and therefore were determined to have a prior knowledge. Others in the crowd had absolutely no experience but was still excited to come up with a guess. Even though a few people came very close to the actual weight, no one got it right. However, when all the answers were aggregated, the result was remarkable. In fact, it was just 1 pound away from the right answer (Galton, 1907). Similar examples such as guessing the number of jelly beans in a jar, have also proved the theory of collective intelligence; namely that by aggregating human judgments from a large, diverse and independent crowd the most accurate answer will appear (Surowiecki, 2004).

Collective Intelligence has been studied in three different categories; *Coordination, cooperation and cognition* (Surowiecki, 2004). Coordination requires a group of people, that all strive for the best solution to coordinate to their best ability. Hence, in order to come up with the best solution "*a person has to think not only about what he believes the right answer is but also what other people think the right answer is"* (Surowiecki, p. 86, 2004). Cooperation refers to how to get people to cooperate even when it seems not to be in their own self-interest, i.e. paying taxes (Surowiecki, 2004) . Hence, cooperation is about overcoming myopic behaviour. The focus of this study is on humans' ability to predict the success of start-ups; thus, the authors focus on the particular branch of collective intelligence; *Cognition*. Cognition focuses on problems that have definitive solutions, such as which start-up is going to generate the highest return of investment and therefore does not take actions of others into account. The experiment to be investigated is based on the work of James Surowiecki, *The Wisdom of Crowds,* that argues if prediction tasks are exposed to a crowd that are *diverse, decentralized and independent* of each other, every perspective regarding the specific event will be covered and hereby the crowd can come up with the most comprehensive solution (Surowiecki, 2004). This study investigates whether certain crowds are more accurate in predicting start-ups than experts in the industry of venture capital. As early-stage start-ups are highly related to uncertainty it is possibly a reliable strategy to depend on human intuition. However, the individual human judgment in decision making is polluted by bounded rationality, which states that humans will always try to act as rational as possible by evaluating all alternative solutions but are limited to do so as there is always lack of information and restricted cognitive ability of humans (Simon, 1987). This often leads to biased interpretation of information and a lack in prediction (Simon, 1987). '*The Wisdom of Crowds'* can act as a solution to bounded rationality (Surowiecki, 2004).

Collective intelligence consists of different characteristics equally to individual intelligence (Woolley, 2010). It is possible to measure collective intelligence and predict the level of how groups perform on different tasks. Collective intelligence performs better than the *"maximum individual intelligence"* (Woolley, 2010). Hence, the performance of groups is argued to be more valuable than the performance of individuals. However, as collective intelligence can be measured and perform differently depending on the task, it is crucial to understand under which circumstances collective intelligence is most effectively exploited. From the results of Malone (2009) it was found that averaging is a good method when uncertainty is present and an estimate of numbers are the target. When averaging is deployed there will occur some random errors which, when not systematically biased but truly random, "*will cancel each other out*" (Malone, 2009 p. 17). Furthermore, biases arise when groups are not diverse enough to cover all the important perspectives of a situation or when early participants affect later ones (Kahneman & Tversky, 1973). Thus, diversity is essential for a group to exploit collective intelligence fully (Surowiecki, 2004; Hong & Page; 2001;2004) This makes it difficult for single hubs or small organizations to exploit collective intelligence effectively as people within a small organization often will be biased or affected by their surroundings, such as the case of a VC.

1.2 The Environment of Venture Capital and the influence of expert knowledge

VC's rely on a few experts when it comes to decision making. Thus, they can be argued to be heavily biased and influenced by each other. For a group to take a wise decision, it requires *diversity* and a portion *of independence* and *decentralization* in order to succeed (Surowiecki, 2004). In the industry of venture capital, a lack of these parameters is highly present. As above mentioned, VCs are in most cases small organizations where people heavily influence each other, and consequently there is a lack of independence. Despite the fact that 45 % of the workforce within VCs are women, only 11 % of them are positioned in high-ranking investment roles (Truong, 2017). *"The lack of diversity (a VC's) influences who gets funded,"* (Truong, 2017). Thus, VCs could benefit from development of higher internal diversity in terms of gender, background, race etc. Diversity in terms of gender and race also affects the investment portfolio of many VCs. 1 % of venture backed start-ups have a black founder, 12 % have an Asian and 87% have white founders (Dober, 2017). In addition, when 60 billion dollars from VCs was disbursed in 2015, only 7% was received by female founders (Dober, 2017). Finally, the essential decisions of which start-ups to invest in are basically dependent on one or a few partners in accordance with top-down management, that acts as a lack of decentralization in the decision processes.

Venture capitalists are considered as experts within industry of VC (Zacharakis & Sherpherd, 2001). Consequently, most studies of VCs have looked into the criteria VCs evaluate start-ups by (Franke et al., 2008; Rea, 1989). The most general evaluation criteria can be divided four categories 1) the product, 2) the team, 3) the market and 4) return on investment (Franke et al., 2008). Franke et al. (2008) showed that venture capitalists evaluate the characteristics of start-ups differently dependent on their prior experience (Franke et al., 2008). Yet this do not show that experienced VCs are more capable at evaluating start-ups' success, than their less experienced colleagues. Actually, Shepherd et al. (2002) found that there exists a curvilinear relation between experience and reliability. Therefore, experience is advantageous to a certain maximum level, but thereafter it becomes problematic due to heuristic shortcuts (Kahneman & Tversky, 1973; Shepherd et al., 2003, Simmons, 2003). Few studies have looked into how VC's can improve their decision-making process. Shepherd & Zacharakis (2003) argues that bootstrapping models, that looks at how VCs reaches a certain prediction rather than looking at the prediction itself, can improve VCs prediction accuracy. Additionally, one of the newest trends within venture capital is the usage of artificial intelligence (AI). In 2009 a company focusing on AI called 'Quid's AI', was asked to use AI in order to predict the future success of start-ups. In the study, the company were asked to come up with a list of 50 companies that the AI technology would suggest becoming the most successful start-ups in the future. In 2016 the 50 listed companies would have become the second best performing VC ever if the portfolio of companies were chosen from the suggestions of the computer (Reddy, 2017). Hence, it is clear that the data-driven approach definitely has its advantages and there is no doubt that venture capital and other industries will become more dependent on big data and AI. Yet the development of machine learning still has its boundaries, as it currently is not able to predict information that is considered "soft information" (Petersen, 2004). Soft information is related to decisions that holds high uncertainty and often needs to be dealt with through intuitive decision making (Petersen, 2004) such as the case of judging start-ups. Intuitive judgments are based on the decisions maker's own knowledge, without him being able to fully explain the underlying factors of these judgments (Simon, 1987). Intuition is closely linked to tacit knowledge (Polanyi, 1966) being the fact that we have more knowledge than we can express. Intuitive knowledge is unique to individuals (Shirley & Langan-Fox, 1996) as they store it over a lifetime of events. Hence, "soft information" and "intuitive decisions" are essential to the investments of venture capitalists. As Andreas Thorstensen, Partner at EQT ventures puts it: "*AI is good for filtering out the noise, but the decision to invest or not will always be about instinct at the end"* (Palmer, 2017).

A two-decade study investigating humans' ability of to foresee the future through forecasting tournaments showed that experts are no better than non-experts at creating predictions (Tetlock, 2005). Additional studies at the World Cup of Soccer in 2002 (Andersson et. al, 2005), and stock prices (Önkal & Muradoglu, 1994) also showed that non-experts are in fact as good as experts when predicting future events. Experts might suffer from an overload of causation as they hold more information to shuffle between (Simon, 1987). This is backed up by Anderson (2005) that proved information effects to play a distractive role in decision making. Experts' lack in their predictions are also likely to be an effect of bias. Overconfidence is a common known bias to experts' decisions and can cause poor decision making (Glaser, Langer, & Martin, 2012). Venture capitalists are believed to be overconfident, due to their low success rate of predicting start-ups. Tetlock (2005) found that forecasters who related everything to a central vision, recognized as experts, were confident in knowledge and predictions despite their predictions being inaccurate. This corresponds with numerous findings in forecasting literature that has proven experts to be overconfident in their predictions (Lawrence et al., 2006) Overconfidence among experts have been shown in i.e. predictions of Russian manager's forecast of future economic (Aukutsionek & Belianin, 2001), and of future recession (Braun & Yaniv, 1992). Hence, the question to be raised is, if there is even a need for experts or non-experts when looking into a field with such a high degree of uncertainty and risk as venture capital. But the answer seems to be yes, as experts and non-experts are found to outperform chance (Anderson et al., 2005; Sundali & Atkins, 1994; Forrest & Simmons, 2000) and experts are seen to be more accurate in their forecast's than statistical models (Armstrong, 1983) Additionally, studies within fields of forecasting in strategic intelligence (Mandel & Barnes, 2014), and forecasting of exchange rates (Önkal et al., 2003) has proven experts to be more accurate in their predictions than non-experts. However, there exists no evidence to whether it is best to have experts or non-experts making the investment decisions within VC.

Collective intelligence has been applied in many different shapes to real-life instances. Most commonly companies have exploited collective intelligence with no markets. An example of this is InnoCentive, a crowdsourcing platform for companies to exploit the minds of a large network to come up with new innovative solutions (Afuah & Tucci, 2012). Additionally, platforms such as Yelp and Tripadvisor have used collective human judgments to rate experiences all over the world. Intrade is a real-life example of how collective intelligence have been exploited through the use of markets. Intrade has predicted future outcome of events such as political campaigns and Box Office rates of upcoming movies since 2001 trough prediction markets and have proven to come up with extremely accurate results (Intrade, 2018). Here individuals' predictions are aggregated and

expressed as “prices” on stocks that can then be traded.
Hence, Collective Intelligence is not an entirely new phenomenon and have been exploited in many real-life settings. However, very few examples exist on exploitation of collective intelligence within the industry of venture capital.

1.3 Research Statement

The objective is to study the phenomenon of collective intelligence and more specifically apply *The Wisdom of Crowds* to investigate whether it is possible to predict the future outcome of start-ups more accurately within the setting of venture capital. As mentioned, the success rate of VC funds is relatively low and has been stated by e.g. Bill Gates to be “pathetic” (Nisen, 2014). New trends within VCs show that data driven strategies are in some stages a reliable source of prediction. Though, early-stage investments are considered to be surrounded by high uncertainty and therefore, decision making in this stage requires a high level of soft information to be interpreted. Prediction with high uncertainty is reliant on human capabilities and their ability to make intuitive decisions. This study seeks to understand whether a large, diverse and independent crowd are better than a few experts at predicting future success of start-ups. The study explores the performance of both internal and external crowds venture capital. Finally, it is studied if overconfidence act as bias to the crowd’s predictions.

Therefore, the research statement for the project is as follows:

How can crowds’ predictions be used to optimize the low success rate of Venture Capital and are predictions victim of overconfidence?

1.4 Methodological approach

To deal with the research question whether a crowd can utilize its collective wisdom and predict the success of start-ups, the authors have constructed a crowd forecasting tournament based on the theory of *The Wisdom of Crowds* (Surowiecki, 2004) As diversity is a crucial factor for the accuracy of *The Wisdom of Crowds*, it is crucial to reach a minimum number of participants. For the diversity of a crowd to exceed the ability of a crowd, the crowd must reach 16-20 participants (Keuschnigg & Ganser, 2017). In addition, it has been proven that increasing the size of a crowd will lead to a higher accuracy (Keuschnigg & Ganser, 2017). To execute the crowd forecasting tournament a quantitative survey was constructed. Here the participants are presented to four historical start-up cases, that have all acted as venture cases. The start-up cases are presented to the crowd in a 2-page PDF-file, and reflects to the researcher's best ability, standard VC evaluation criteria. The VC has exited on all the cases. In other words, they are no longer involved with the start-ups. As the four cases are historical, their respectively return of investment is known to the VC and the researcher. All of the four cases have given different return of investments, more specifically 10 x-,8 x-, 3 x- and 0 x return of investment. Two of the start-ups are considered by the researchers to be "good investments" and two being "bad investments". After being presented to the four cases, the participants are asked to assess and invest 1000 points in the four start-ups in order to make the investment that generates the highest return. The main objective for the participants is to identify which of the four start-ups that has provided the best return of investment. As the researchers are not that interested in the individual's predictions, the participants pooled into crowds. The composition of the crowds is to be explained further later on in the study. However, the most essential crowd division will consist of *experts* and *non-experts.* Experts are determined to have working-experience within VCs. The expert crowd gathered from both Copenhagen and London, consist of 7 experts, which is a suitable number for a knowledgeable crowd (Mannes, et al., 2014). The non-expert crowd consists of 58 non-experts. Previous research has proven how a larger group of non-experts, in combination with a small group of experts are capable of predicting uncertain future outcomes (Surowiecki, 2004, Galton, 1906) Hence, the crowds' predictions will be analysed - both in combination of each other and separately.

1.5 Implications

The purpose of the study is to contribute with new knowledge of whether collective intelligence have a role to play as a reliant predictor of investments within Venture Capital. The hope is to optimize decision making in venture capital funds and improve the success rate of investments. This study could eventually, dependent on the results, contribute to the development of a new business model within venture capital. Though, the study holds some limitations. The cases are presented by the research team and are not the original material the investment was based on. Furthermore, investments are made years back where the investment landscape can have had different characteristics, though no cases are from later than 2011, so this should not play a large factor. Finally, VC's are experts within their field and could have prior knowledge to some of the cases, which would hold a huge limitation. Though, this should have been avoided as all the cases are foreign and VC's were told not to answer if they knew the cases prior to the presentation and this does not seem to be reflected in the results. Experts in this study are Venture Capitalists' from leading VC's in Copenhagen and London.

Chapter 2 - Structure

The following section will introduce the overall structure of the thesis. **Chapter 3** is the literature review that will explore the fields of *Collective Intelligence*, *Venture Capital* and *Forecasting*. Hence, this chapter contains the theoretical part of that this study builds on. The part that describes Collective Intelligence will present the underlying structures that have led to the use of Collective Intelligence. Additionally, Collective Intelligence will be investigated in the setting of groups that interact and groups that do not interact. Finally, will *The Wisdom of Crowds* which is a variety of Collective Intelligence and act as the main theory of this thesis, be explored.

The part that describes venture capital, will mainly focus on the criteria venture capitalists evaluate start-ups by and how they in general evaluate. Furthermore, the part will introduce how venture capitalists perceive success. Finally, an international perspective of how cultural differences and ecosystems affect the performance of VC's, will be explored.

The part of forecasting will introduce theory on how studies of forecasting have been investigated within other fields and especially look into how expert- vs. non-expert crowds have been studied within forecasting. Additionally, it will be explored how calibration and overconfidence have been studied within other fields. Finally, this part will investigate how venture capitalists

Chapter 4 is the hypothesis development. Here six hypotheses will be developed based on the introduced literature in order to answer the research question. The hypothesis will roughly investigate three areas of crowd prediction. First it will be compare the crowd predictions of non-experts and experts separately and combined. Secondly, it will compare crowd predictions of experts with certain characteristics. Thirdly, it will be investigated whether bias act as an obstacle to the predictions and be compared if the different crowds hold different bias to their predictions.

Chapter 5 describes the methodology of the study. Hence, this chapter describes the choices taken by the researchers in order to investigate the subject at matter in the best possible way. This ranges from i.e. research approach to data handling. Moreover, it sets to introduce the start-up cases that have been used to conduct the predictions. The chapter will briefly explain how the cases were obtained, and finally, each start-up will be presented from a perspective of *the product* and *the team*.

Chapter 6 consists of the findings from the quantitative crowd prediction tournament. First, the most important and overall findings are presented. Secondly, the findings of each hypothesis are presented.

Chapter 7 is the discussion and will be divided into two parts. The first part will discuss directly on the findings of the six hypotheses. In the second part, a general discussion of the findings will be elaborated on.

Chapter 8 is the conclusion where the researchers will to their best ability try to give an answer to the research question. This will be done by summarizing the findings of the six hypotheses.

Chapter 9 will elaborate on the limitations of this study and discuss how the quality of the study could be enhanced seen in retrospect.

The purpose of **Chapter 10** is to give an understanding of how future studies could be conducted and encourage other researcher to look into the field of prediction within venture capital.

Finally, **Chapter 11** will introduce the practical implications that the findings have for venture capital.

Chapter 3 - Literature Review

The following section will introduce the literature that is needed in order to answer the research question. The purpose is to frame theories and previous studies making it possible to test the six hypotheses. Three overall concepts are investigated in the literature review; *Collective Intelligence, Venture Capitals* and *Forecasting* which will be covered in that particular order.

3.1 Bounded rationality

Rationality is in economics referred to as an actor seeking to maximize value given all information (Jones, 1999). Decision making, in reality, often strives to be as rational as possible but are always limited to some extent, which leads to the decisions being irrational (Jones, 1999). Huymans (1970) has identified three explanations for why irrationality occurs. First, decision makers have limited cognitive ability. Secondly, humans by nature is not rational (Kahneman & Tversky, 1973) and thirdly, the amount of information is either too little or too large and complex, making it impossible for humans to take rational decisions (Brunsson, 1982).

In alignment, Kahneman (2003) introduced the idea of thinking fast and slow. He argues that individuals have two systems that controls the way of thinking. System1 is the fast system that generates conclusions based on i.e. intuition, emotions etc. that is automated and generates conclusions (Kahneman, 2003). System2 is effortful and slow thinking system (Kahneman, 2003). When System2 is deployed it is likely to be a more controlled way of thinking than that of System1. As "*the overall capacity for mental effort is limited"* (Kahneman, 2003, p. 698) Kahneman argues system 1 are to take over when system 2 is actual in need for.

Herbert A. Simon (1987) introduced this as the term "Bounded rationality". Gigerenzer & Selten (2002) argued that Bounded rationality is when humans strive to act as rational as possible but are restricted to do so, as limited information is present or there is a lack of cognitive ability (Jones, 1999; Huymans, 1970) Bounded rationality occurs due to an overload of impressions from a variety of sources such as intuition, accessibility of information and knowledge, prior experience and the ability to frame a problem (Kahneman, 2003). Simon (1955) describes in his early work that decision making is a process steered by aspiration levels. Aspiration can be explained as the satisfactory alternative when an individual is in the process of decision making. If the alternative is difficult to find, a persons' aspiration level might decrease. On the other hand, if alternatives are easy to find an if these are considered to be at a satisfactory level, the aspiration level will rise (Simon, 1987). This leads to individuals taking mental shortcuts also known as heuristics. As individuals have a limited mental capacity, Collective Intelligence can act as a solution to bounded rationality (Surowiecki, 2004).

3.2 Collective Intelligence

3.2.1 Origin

Collective Intelligence is a term that has significantly spread with the digital era. Though, the idea of Collective Intelligence can be tracked all the way back to Aristoteles and his work from the antiquity (Waldron, 1995)**.** Collective Intelligence has a long history in psychology (Malone, 2009) Moreover, William Wheeler (1911) has described social insects and its intelligence as a "superorganism". A swarm of insects is collectively capable of solving complex task such as nest building or route optimization (Bonabeu & Meyer, 2001).

The concept of collective intelligence is first introduced by Pierre Lévy (1997). He describes it as a form of universal distributed intelligence which is in real-time continuously improved and coordinated. This will ultimately result in an effective utilization of capabilities between individuals (Lévy 1997). The most recent research has shown that crowds have a collective intelligence that can provide solutions, that are very innovative and possibly able to forecast future events very precisely (Hong & Page, 2001; 2004; Surowiecki, 2004). The most recent and commonly known way of utilizing collective intelligence is in politics (Landemore, 2008; Surowiecki, 2004). Additionally, companies such as Google, Affinova, InnoCentive have all successfully managed to use collective intelligence in e.g. R&D, knowledge management and customer service (Afuah & Tucci, 2012). This task is typically referred to as Crowdsourcing. The first real usage of the term crowdsourcing was invented by Jeff Howe in 2006 (Howe, 2006). It originally refers to the task of outsourcing a specific job to an external crowd in order to harness the benefits of collective intelligence (Howe, 2006).

3.2.2 Collective Intelligence

Woolley (2010) investigated whether a collective intelligence existed and if it has different characteristics than the intelligence of individuals. He found that groups' intelligence could be measured and that the level of the groups' performance on different tasks could be predicted. The results showed that collective intelligence is present and that collective intelligence perform better than the *"maximum individual intelligence"* (Woolley, 2010).

Collective Intelligence can be exploited through group work. In fact, groups are shown to perform better than individuals even when the group members are not believed to be very knowledgeable (Surowiecki, 2004). Actually, large groups consisting of less informed members, can still reach collective wise decisions (Surowiecki, 2004). In fact, studies at the World Cup of Soccer in 2002 (Andersson et. al, 2005) and forecasting within political judgement (Tetlock, 2005) have shown non-experts to reach better outcomes than experts. Additionally, it is proven that adding diversity to a group leads to more accurate decisions than adding ability to a group (Keuschnigg & Ganser, 2017). In other words, it is more valuable to add an extra individual with a different point of view than the rest of the group members to a group, than it is to add an extra individual who is knowledgeable of the specific field investigated, if this person doesn't add to an increase in the diversity. Additionally, Keuschnigg & Ganser, (2017) found that having as many people in the judgement process as possible, will lead to higher collective accuracy. However, it is required that there is no systematic bias - only a random error (Malone, 2009).

Malone (2009) investigated how collective intelligence in groups is used most effectively. He defined Collective Intelligence *"as groups of individuals doing things collectively that seem intelligent" (pp.2)*. From the results, it was found that averaging is a good method when uncertainty is present and an estimate of numbers are the goal (Malone, 2009). When averaging is used there will be some random errors which, when not systematically biased but truly random, will cancel out each other (Malone, 2009). Larrick and Soll (2006) also argues that many people fail to exploit the advantages and the power of averaging, as many individuals tend to think of averaging as creating an answer of average quality. In addition to this, many people tend to place too much emphasize in expert opinions and rely on single expert's as sources of information (Larrick & Soll, 2006). Hence, Larrick & Soll (2006) argues averaging results from a crowd as a reliable tool for collective accuracy.

The accuracy of Collective Intelligence has been studied for many years. Most research, however, is related to interaction patterns and communication in group work (McGrath, 1984; Steiner, 1972). In other words, when people in a crowd are cooperating with each other in order to come up with the best solution to a given task. Shaw (1932) found in his experiment of problem-solving tasks, when comparing groups and individuals, that 53% of groups managed to come up with the right answer compared to only 7.9% of the individuals. Collective accuracy can be a product of group work but the results of groupwork can also end up as a product of bias (Alper et al., 1998). Hence, group work can lead to biased decisions (Alper et al., 1998). In some situations, groups might benefit from interaction and discussion, however, too much communication might lead to the group being less intelligent (Surowiecki, 2004), due to biases such as groupthink or lack of diversity. It is argued that biases arise when groups aren't diverse enough to cover all important perspectives or when early participants affect later one's biased decisions can occur (Malone, 2009). This could to some extent be solved by the use of large groups, however, large groups are difficult to manage and therefore can become inefficient (Keuschnigg & Ganser, 2017; Surowiecki, 2004).

Thus, researchers have investigated the efficiency of collective intelligence through groups that are independent of each other and not able to interact (Surowiecki, 2004; Hong & Page, 2001; 2004; Galton, 1906). Results have found that the collective intelligence of independent groups that are not able to interact exceeds the intelligence of individuals (Surowiecki, 2004; Hong & Page, 2001; 2004; Galton, 1906). Galton (1906) did an experiment on collective accuracy. At a county fair, he invited a large number of people to a weight guessing competition of an Ox (Galton, 1906). Around 800 people participated – here the participants were both experts and non-experts. The one who came up with the most accurate guess would win a price. Galton analysed the answers from all participants and discovered that not a single person got it right. More interesting, he found that the average result of all the guesses was just 1 pound away from the real weight.

James Surowiecki (2004) argues that there is no limit for the complexity of the task that a crowd are able to handle. Though, he argues that collective intelligence can be divided into three types.; Cognition, Coordination and Cooperation. Figure 1 a model of Collective Intelligence that is based on the work of Surowiecki (2004).

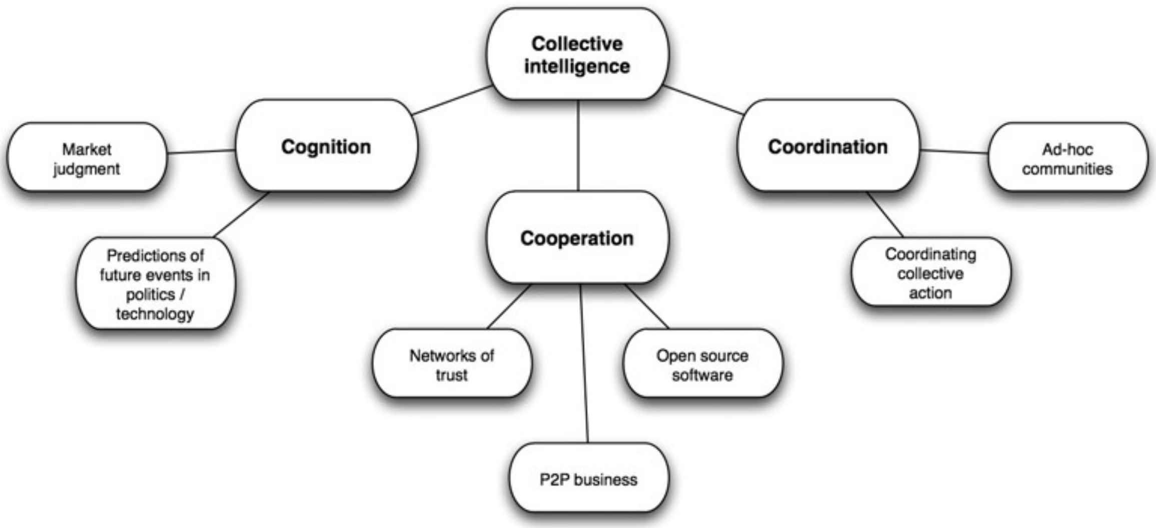


Figure 1 - Source: Generozova, 2006

Cognition is when the task or the question that are being asked to the crowd, has a definitive solution, e.g. *"How likely is it that this drug will be approved by the FDA?"* (Surowiecki, 2004, p. 26). One way of dealing with this problem is to have diversity and independence. When dealing with cognition problems, it is key that individuals do not modify their answers in order to satisfy everyone but to come up with a solution that pleases themselves. Brenner, Griffin, and Koehler (2005) found that groups generated more optimistic predictions than those of individuals. Therefore, it is crucial that members think and act as independently as possible. One example is when a US submarine, The Scorpion, disappeared in 1968 and was to found again. The only cue known, was that the submarine disappeared on its way from the North Atlantic to Newport. Instead of consulting with a few submarine and ocean experts, it was decided to ask a large number of individuals, ranging from mathematicians, salvage men and submarine experts. They were asked to come up with their best guesses independently of each other. Despite the almost impossible task, their aggregated answer managed to locate the lost submarine (Surowiecki, 2004).

Coordination relates to problems where members of the crowd must coordinate their behaviour with each other knowing that everybody else are doing the same. E.g. "*How do you drive safely in heavy traffic?"* (Surowiecki, 2004, p. 27) In other words, a problem where the individual has to think not only on what he or she believes is the right answer but also have to take others people's thoughts into account. Therefore, many forms of coordination problems require bottom-up solutions for people in groups to voluntarily coordinate their actions in an efficient way (Surowiecki, 2004).

C*ooperation* involves the challenge to make distrustful people work together despite individuals pursuing self-interest, such as paying taxes or dealing with pollution. The coordination problems have its similarities to coordination problems, as people need to think about others opinion in order to come up with a good solution. However, in cooperation problems, it is crucial to establish trust within the members of the group in order to avoid myopic behaviour (Surowiecki, 2004). A political scientist Robert Axelrod argued that cooperation is not only about trust but the durability of the relationship (Axelrod & Douglas, 1988). When people interact and repeatedly trust each other, there will be an incentive to cooperate and therefore they will not take advantage of each other (Axelrod & Douglas, 1988).

3.2.3 Prediction

This thesis strives to test collective intelligence in a complex setting of predicting future events - namely the success of startups. Therefore, the focus will be on *cognition tasks* and more specifically, *predictions* through *The Wisdom of Crowds*. A crowd is argued by Le Bon (1995) to be more than just the sum of its members. Instead it is an independent mechanism that has the will of its own. It often acts in ways that no one in the crowd originally intended. Crowds are able to come up with intelligent solutions in various settings given the right circumstances (Surowiecki, 2004, Le Bon, 1995). There are different methods of using crowds as a source of prediction, which can be either tournament based or a cooperative based (Afuah & Tucci, 2012). When used cooperatively the individuals in the crowd are able to discuss and provide inputs on alternative ideas to finally come up with the best solution. Alternatively, tournaments can be executed, where the participants in the crowd are not cooperating but competing to provide the best solution.

3.2.4 Wisdom of the Crowd

"The best collective decisions are the product of disagreement and contest, not consensus or compromise" (Surowiecki, 2004, p. 28). According to Surowiecki (2004), the collective wisdom of a group surpasses the smartest individual member. People tend to lack the ability to see into the future as predicting a future outcome is highly correlated with uncertainty and risk. Most individuals struggle with making simple and sophisticated cost-benefit calculation and instead let emotions affect

their judgment, which leads to biased and suboptimal decisions (Surowiecki, 2004; Kahneman, 2003). Despite these limitations for individuals to forecast precisely, the aggregation of all the imperfect judgements will eventually lead to our collective intelligence being excellent (Surowiecki, 2004; Larrick & Soll, 2006; Keuschnigg & Ganser, 2017). When combining forecasts, the information stream increases as individuals would have retained their information from different sources leading to a more accurate prediction (Lawrence et al., 2006). This intelligence is what Surowiecki (2004) calls 'The wisdom of crowds'. He has identified three necessary conditions in order for a crowd to be wise: *Diversity, Independence and decentralization.*

Diversity is when individuals of the crowd is considered to have different point of views to a given task in order to achieve the best possible solution (Hong & Page, 2001;2004, Surowiecki, 2004). Different point of views is often reflected on the individuals background, hence, diversity can refer *"to differences in demographic characteristics, cultural identities and ethnicity, and training and expertise"* (Hong & Page, 2004, p1). Contradictory, if crowds possess the same point of views and has the same amount of information, the answer or solution will not cover all perspectives (Hong & Page, 2004). When e.g. building a startup or investing in one, diversity is important because it will give meaningful differences of ideas and attitudes instead of minor variations around the same concept (Surowiecki, 2004).

Independence is, explained in short, when the opinions of a participant is not influenced by others in the crowd. James Surowiecki (2004) argues that the best collective decisions are a result of disagreement and contest instead of consensus and compromise. In order to obtain the best solutions, members are not asked to modify their answer in order to reach something that everyone can agree on. Instead, by aggregating the human judgements in a way that represent what everyone thinks (Surowiecki, 2004; Larrick & Soll, 2006). Despite, the general rule of thumb of small and coordinated groups being able to produce better products than the individual, James Surowiecki (2004) argues that the more independent the group is the more intelligent. Dependence among crowd members are found to be negatively correlated with crowd performance (Hong et al., 2016)

Decentralization is when decisions are not centralized around few people (Surowiecki, 2004). A good example is Linux, an open-source system, that has no formal organization behind it but instead driven by external contributors from around the world. Decentralization leads to diversity in opinion and decreases bias as it distributes knowledge and decisions to a variety of individuals (Surowiecki, 2004).

Most studies of the *The Wisdom of Crowds* have been conducted asking individuals to guess a single magnitude estimate (Galton, 1906, Larrick & Soll, 2006). However, in less tested areas, the effect of The Wisdom of Crowds has also proven to work in more complicated settings where the participants were asked questions regarding rank-ordering tasks (Lee et al., 2012; Grazing 2016). In the research of Gordon (1924), he asked individuals to rank the weights of Ox's, from the heaviest to the lightest. The results of this showed that when he increased the group size, the correlation between the true values and the collective judgments followed (Gordon, 1924). Additionally, a group of California scientist plunked 20$ on a horse race (Grazing, 2016)) The goal was to pick the first-, second-, third- and fourth-place finishers. The odds of being correct were obviously not good, however, the potential outcome was high. They asked a small group of experts within horse racing and afterwards pooled their insights and knowledge to come up with an answer. Very interesting, none of the horse enthusiasts who took part in the experiment was right about the combination of horses, nor was any of the race track's own experts but by combining the expertise and tapping into The Wisdom of Crowds of all the participants at the horserace managed to get the superfecta – the right ranking of horses (Grazing, 2016).

The two sociologists Jack B. Soil and Richard Larrick (2016) has stated that individuals need to stop *'chasing the expert'*. In many industries people rely on few leading experts, trusting them to solve a certain problem and to take the right decision (Surowiecki, 2004). Even when a crowd, consisting of a majority of non-experts, predict the outcome of an event, say a horse races, most people assume that putting their money and trust on few individuals is the right decision (Larrick & Soll, 2006). James Surowiecki (2004) believes this is a mistake and his theory 'The Wisdom of crowds' strive to act as a solution to this specific problem.

3.3 The Industry of Venture Capital

3.3.1 Definition of Success:

In order to understand how venture capitalist's think, it is important to understand how they do business. Even though, venture backed startups have a higher success rate than startups that are not venture backed (Rajan, 2010), the failure rate of investments in the industry of Venture Capital are relatively high (Lemkin, 2012; Appendix 7). Though, in order to understand the failure rate, it is essential to know what is considered as a success. In venture capital, there is a rule of thumb called the 10x-rule or the "power law" (Lemkin, 2012). The 10x-rule is an expression of the expected return of investment when Venture Capitals (VC's) invest in startups. Although a 10x return of investment could seem high, this is the expected return as VC's argue that the risk and uncertainty of investing in early-stage startups is extremely high. Therefore, they only expect one or two out of ten startups to give the expected return of investment (Lemkin, 2012; Appendix 7). Or as Co-founder and CEO of Beco Capital puts it "*our business model is governed by the "power law:" what that means in essence, is that out of every ten early-stage investments, around two will create all the returns and the rest will underperform by generating little to no returns"* (Griffith, 2014). VC's tend to invest in less than 1% of the investment deals that they see (Appendix 7). Hence, the likelihood of picking the best performing deals is a difficult and tough process.

3.3.2 The International Environment – A comparison to London

Most job creation and economic growth within a country is established from fostering new companies that then develops into larger companies (MandagMorgen, 2015). A new key concept within the industry of venture capital, called 'Scaleups' is defined as companies who, annually, have an average growth of employees or turnover by at least 20% in three years (MandagMorgen, 2015). Denmark is among the nations that produce the largest number of startups pr. citizen (KPMG, 2017; MandagMorgen, 2015) and has created an accessible environment for entrepreneurs to launch new ideas. However, none of these, are candidates of being a new multinational corporation like Novo Nordisk or Lego (MandagMorgen, 2015). So how do Denmark, as a nation, foster more startups and support them into even more growth and ultimately being a multinational corporation? It is a combination of innovative entrepreneurs, risk seeking investors, experienced advisors and lastly, a business environment capable of providing a workforce and technology (MandagMorgen, 2015). This is, in other words, called the '*Ecosystem'*. Silicon Valley, New York and London are known as having the strongest and most healthy Ecosystem of fostering both startups and scaleups (StartupGenome, 2017)

With 4,300-5,900 active startups, London has the 4th highest number of start-ups in the World and is argued to possess the strongest Ecosystem in Europe (StartupGenome, 2017). With its powerful financial capabilities and an impressive ranking among startup experience, improvements in funding, performance, market reach and talent - London has created an Ecosystem where start-ups have the ideal surroundings for further growth (StartupGenome, 2017). Denmark, on the contrary, whose

Ecosystem is mainly focused upon Copenhagen and Aarhus is ranked as the second best Ecosystem in the Nordic region, marginally behind Sweden. Copenhagen is considered by experts to be a large player within early-stage start-ups but lack the ability to maintain growth within its boundaries (KPMG, 2017). Other Ecosystems, such as London, Silicon Valley or even Berlin are attractive to entrepreneurs to foster growth at a later stage (StartupGenome, 2017). London in particular has a strong Ecosystem due to powerful tech-companies, high-level banks and strong VC's. Therefore, start-ups have great access to investors and potential acquires (Ciszewski, 2016).

Psychologist Dr. Geert Hofstede, famous for his work on cultural differences (Hofstede, 2011), has in his studies created a recognized and international standard for identifying cultural differences. The model of Hofstede (2011) consists of 6 elements;

- *Power distance* which is related to the fact that all individuals are unequal
- *Individualism* is regarding the independence in the society
- *Masculinity* is related to the motivation and desires in life.
- *Uncertainty Avoidance* which is the level of stress in a society in the face of an unknown future.
- *Long term Orientation* which is the choice of focus for people's effort in terms of the future, present and past.
- *Indulgence* is how individuals try to control desires and impulses

Hofstede (2001) argues that nations are not the best way for studying cultures - however, they are usually the only source available for comparison. Other studies have shown that factors such as legal rights and protection, the stock market development and cultural differences has a strong impact of VC's success (Nahata, 2014). The study also showed that mutual trust between the VC and start-up is an important factor which is also correlated with the dimensions of Hofstede (Chakrabarti et al. 2009). Thus, VC's are more likely to do a thorough investigation when screening start-ups where cultural differences are expected to be present (Nahata, 2014).

3.3.3 The investment process of VCF's:

The investment process of VC is a rather long and complicated process. Tyebjee & Bruno (1984) developed a five-stage model in order to explain the activity of VC's investments. The model, Figure 2 below, describes these steps as an orderly and chronological process. The five steps consist of 1) deal origination, 2) deal screening, 3) deal evaluation, 4) deal structuring and 5) post-investment activities.

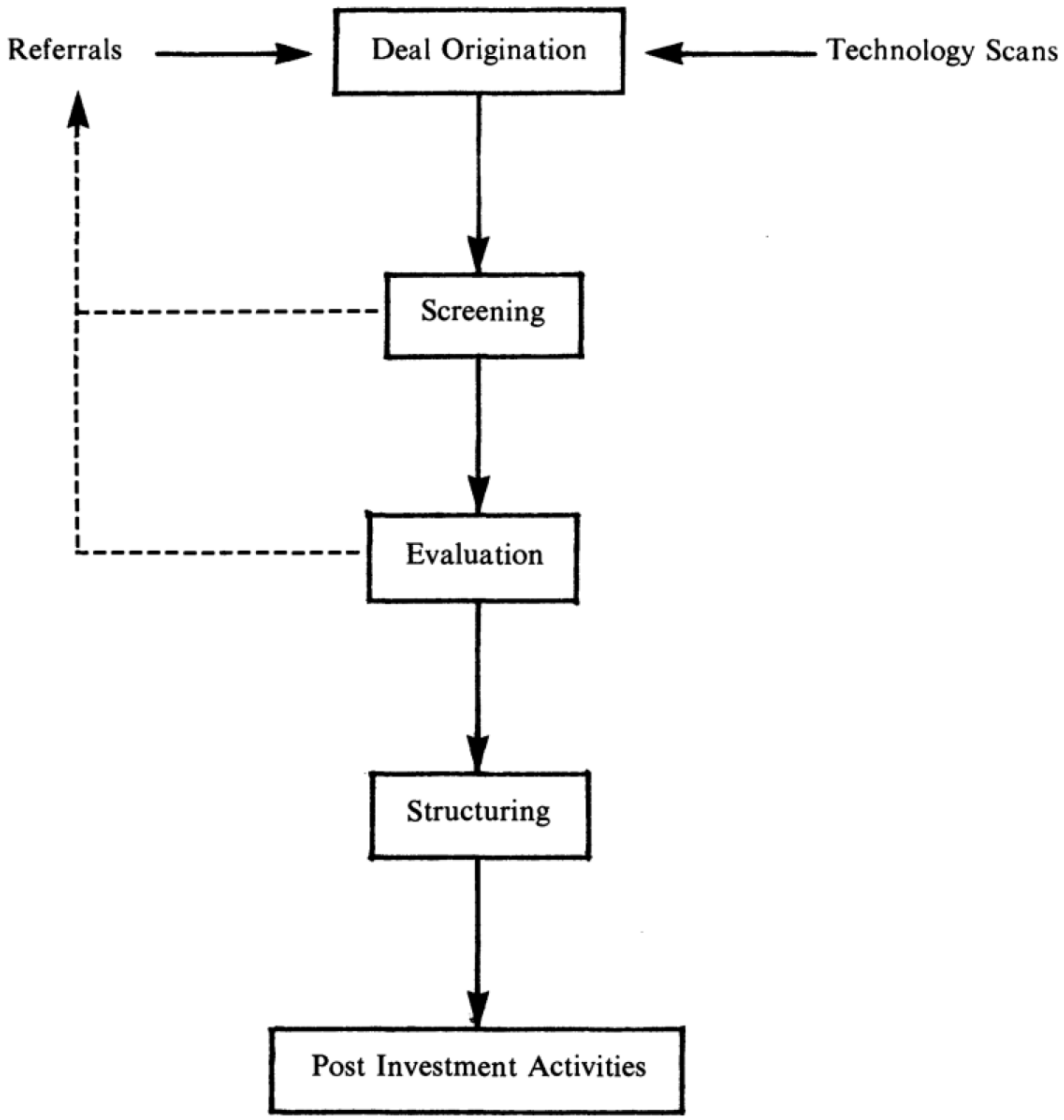


Figure 2 - Source: (Tyebjee, 1984)

Deal origination describes the process when VC's become aware of potential investments. This process commonly occurs through referrals, start-ups that addresses the VC's or vice a versa. Some newer examples have used Artificial Intelligence (AI) to become aware of potential investments

(O'Hear, 2017). The use of AI is a new trend within VC's and is used to screen investment cases and find these through the internet. This is an increasing tendency within venture capital.

Screening describes the process of VC's screening the many investment deals they are faced with, in order to eventually find the potential best investment cases. Normally, VC's invest in less than 1% of the cases presented to them which indicate the heavy workload related to this specific task (Farha, 2016).

Evaluation describes the process of VC's assessing deals based on certain evaluation criteria. This process involves a great deal of subjectivity and gut feeling of the venture capitalists. The subjectivity is highly involved as information symmetry is present due to high amount of risk involved in assessing start-ups, as their future are unknown. The evaluation stage is an essential phase, as this is when venture capitalists judge good- and bad investments from each other (Tyebjee, 1984). The risk and evaluation criteria depend of the maturity of the start-up.

Deal structuring describes the process where the two counterparts, consisting of the VC and the start-ups, have to formulate a mutually accepted agreement, in order for the VC to invest. Setting up an investment has different purposes. First and foremost, it focuses on a price that the VC's should buy a share of equity for. Secondly, are agreements of ground rules recognized in order to prevent high management salaries, limit capital expenses and establishes board structure etc. (Tyebjee & Bruno, 1984)

Post-investment activities describe the process of how the venture capital interact after having agreed on a deal with the specific start-ups. In this process, the focus of the VC's changes from being an investor to being a partner. Thus, the interaction becomes more cooperative than evaluating (Thybjee & Bruno, 1984). The level of interaction between VC's and the start-ups differ from deal to deal, as it depends on the strategy of the specific VC. Commonly, VC's try to have an external position in regard to the day-to-day practice and advice mainly on strategic and operational issues (Appendix 7.1). Research have shown that start-up backed companies that have a VC within their board are more likely to have success in the long run (Guo, 2015). Hence, some interaction between the VC and the start-up seems to have positive effect on the outcome of the investment.

This study focuses on the screening and evaluation process of Tyebjee & Bruno's (1984) five stage decision process model of venture capitalist investment activity. This focus is chosen, as the screening and evaluation stage is where subjective considerations and critical investment decisions are made.

3.3.4 Venture evaluation criteria

Since the 1970's the interest of evaluation criteria venture capitalist's asses their investments by have been of great interest to academics and a topic of much research (Franke et al., 2006; Wells, 1974). According to Franke (2008) the strong interest seems to be explained by three reasons; first, that it is of great interest to entrepreneurs as it can help them when seeking funding. Secondly, it helps VC's to compare their findings to other experts. Finally, the criteria are believed to be success factors for startups, as VC's are assumed to be experts. The criteria venture capitalists evaluate companies by have been studied from many different angles and mentioned as many different names. Figure 3 below categorizes the most important evaluation criteria found in the existing venture capital literature over time.

Figure 3 – Source: Franke et al. 2008

Survey of the Literature

Author(s)	Sample	Method	Evaluation criteria by rank order of importance
Wells (1974)	8 VCs	Personal interviews	(1) Management commitment (2) Product (3) Market
Poindexter (1976)	97 VCs	Mail survey	(1) Quality of management (2) Expected rate of return (3) Expected risk
Johnson (1979)	49 VCs	Mail survey	(1) Management (2) Policy/strategy (3) Financial criteria
Tyebjee and Bruno (1981)	46 VCs	Phone interviews	(1) Management skills and history (2) Market size/growth (3) Rate of return
MacMillan et al. (1985)	102 VCs	Mail survey	(1) Capability for sustained intense effort (2) Familiarity with the target market (3) Expected rate of return
Goslin and Barge (1986)	30 VCs	Mail survey	(1) Management experience (2) Marketing experience (3) Complementary skills in team
Robinson (1987)	53 VCs	Mail survey	(1) Personal motivation (2) Organizational/managerial skills (3) Executive/managerial experience
Rea (1989)	18 VCs	Mail survey	(1) Market (2) Product (3) Team credibility
Dixon (1991)	30 VCs	Personal interviews	(1) Managerial experience in the sector (2) Market sector (3) Marketing skills of management team
Muzyka et al. (1996)	73 VCs	Personal, standardized interviews	(1) Leadership potential of lead entrepreneur (2) Leadership potential of management team (3) Recognized industry expertise in team
Bachher and Guild (1996)	40 VCs	Personal interviews	(1) General characteristics of the entrepreneur(s) (2) Target market (3) Offering (product/service)
Shrader, Steier, McDougall, and Oviatt (1997)	214 new ventures with IPO	Interviews, publicly available documents	(1) Technical education (2) New venture experience (3) Focus strategy
Shepherd (1999)	66 VCs	Conjoint experiment (personal/mail)	(1) Industry-related competence (2) Educational capability (3) Competitive rivalry

As seen many different criteria in many different shapes and called by different names have been studied. However, the criteria venture capitalists evaluate start-ups by differ depending on what stage of maturity the start-ups are in. Bachher & Guild (1996), who investigated evaluation criteria of early-stage start-ups, which is also the focus of this study, found that *"General characteristics of the entrepreneur(s), target market, and Offering (product/service)"* are the most essential to evaluate start-ups by. These findings correspond well with other literature of evaluation criteria of early-stage start-ups (Rea, 1989) and the interviews of Danish investment managers within venture capital conducted by the researchers (Appendix 7). Hence, general evaluation criteria of early-stage start-ups can be divided into three categories – *team, product, and market.* Therefore, these are the criteria that will be explored in the following section.

The Team

Many studies have been conducted on the team behind the startup as it is one of the evaluation criteria that have gained the most attention (Siegel et al., 1993; Bernstein et al., 2016; Franke et al. 2008; Appendix 7). In fact, results show that the most successful and experienced angel investors react primarily to team information (Bernstein et al., 2016). Additionally, Siegel et al. (1993) found that out of the six evaluation criteria he had distinguished between in his study, the team- and the entrepreneur's characteristics were the most valuable, when differentiating between the successful and the unsuccessful start-ups. Muzyka et al. (1996) found that when assessing venture cases VC's believed the *leadership potential of the entrepreneur, the leadership capabilities of the management team* and the *industry experience within the team,* to be the most important evaluation criteria. Franke & Gruber (2008) found that the most important evaluation characteristics of a team are *industry experience*, *educational background* and *leadership experience* but as long as these criteria are present within the team, not all of them have to have industry- and leadership experience, as they found that it was sufficient if only some of the team members held these capabilities. Cooper & Bruno (1977) showed that successful startups are often founded by at least two founders, as 80% of successful high-tech companies have more than one founder (Cooper & Bruno, 1977). Furthermore, heterogeneous teams are by far preferred over homogeneous teams when it comes to educational backgrounds (Franke & Gruber, 2008; Appendix 7).

The Product/firm

The idea and the product offering are vital parts of succeeding with a start-up. Though, if ideas and products are innovative these can be hard to evaluate on (Bruton & Rubanik, 2002). However, the greater the innovativeness of a product the greater is the performance of the start-up (Bruton & Rubanik, 2002). This is a reasonable statement as start-ups flourish from innovation and by setting themselves apart from existing products. Moreover, the greater volume of innovativeness suggests a higher opportunity of obtaining a first-mover opportunity and a competitive advantage compared to other companies, which is a highly valuated criterion by VC's (Imamuddin, 2009). This is also reflected, as an indicator of product value that VC's focus on, is whether a patent is present (Song et. Al., 2008). Companies in possession of unique technology tend to have more innovative products than those of their competitors (Siegel et al., 1993). Moreover, companies with unique technologies are often capable of avoiding competition during the first couple of years, which significantly reduces the competitive risk (MacMillan, et al.,1985) and increases the likelihood of success. Additionally, start-ups with fully developed prototypes are more likely to escape development failures, which is why this is an indicator of a high value to VC's (Macmillan et al., 1985). Another factor that enhance start-ups' possibility of success is their internal assets. Companies with more assets than those of their competitors seems to more successful (Fredriksen et al., 1989) as these companies are more able to deal with unforeseen events and occurrences. Thus, it puts them in a situation, where they have a higher organizational slack.

The market

The market potential is an essential evaluation criterion to VC's for many reasons. VC's objective of an investment is to gain a 10x return of investment within approximately 5-7 years. Hence, the start-ups market potential has to be of great volume in order to increase the possibility for VC's to reach their investment objective (Bruton & Rubanik, 2002). Thus, the possibility of a high growth rate within the market is a vital evaluation criterion for VC's. Additionally, it is of great value to start-ups to have obtained knowledge of the industry in which the company competes (Appendix 7.1). Furthermore, companies in possession of unique technology tend to have a higher opportunity of success in a given industry. Moreover, companies with unique products tends to be supply based and are therefore not dependent of others', which puts them in a greater position of gaining success in the market. Additionally, the customer acquisition cost (CAC), which might differ from the industry of the start-up, is of great importance (Clark, 2016). The total CAC is an expression of how much it cost to gain an extra customer. Though, this number is often relatively fluent in early-stage start-ups and not easy to subtract.

3.3.5 Evaluation vs. experience:

As the assumption is that VC's are experts, it is important to understand what makes a good expert. Experienced experts evaluate differently from less experienced experts (Franke et al, 2008). Even though both experienced and less experienced VC's tend to prioritize *industry experience* and *field of education* among the three most important evaluation criteria, they also differ within evaluation criteria, with the most noticeable being the value that experienced experts assign to *team interconnection*, whereas non-expert tends to focus more on the *individual qualities of team members* (Franke et. al., 2008). With this in mind, common knowledge would suggest that venture capitalists with the most experience are to asses' future investments better than those with less experience, as experienced decision makers tend to exploit decisions processes better than those of non-experts (Anderson, 1983; Dreyfus & Dreyfus, 1986). Though, Sherpherd et. Al. (2003) found that there exist an optimum of experience, as their findings showed that reliability and experience had a curvilinear relation, meaning that experience to a certain level is useful, but then it becomes an obstacle. If the level of experience is too low an information overload can occur that decreases the reliability, but if the level of experience becomes too high a negative relationship to reliability occurs as well (Shepherd et al., 2003). This can be explained through the fact that experienced "experts" tend to use heuristics/mental shortcuts (Simmons, et al. 2003) and thereby use system1 when system2 is needed (Kahneman, 2003). Additionally, experience can result in routed ways of thinking, hence experts becomes affected by their past experiences and misses to recognize new approaches (Tversky and Kahneman, 1973). Furthermore, decision makers with a high level of experience are likely to suffer from overconfidence and being convinced that certain events are to occur, because they have seen them occur before (Mahajan, 1992) Fourthly, decision makers tend to judge and generalize based on small samples, hence concluding on a too narrow foundation (Mahajan, 1992) In other words, if they have seen it happen once or twice before, they are likely to believe it to happen again, even though a sample of one or two cases are not nearly enough to make it significant.

3.4 Forecasting

However, studies within other fields have shown that experience within forecasting is a clearly defined advantage (Mandel & Barnes, 2014; Keren, 1987). Mandel and Barnes (2014) investigated the accuracy of forecasting within strategic intelligence. Strategic intelligence is used to assists decisions makers such as government leaders in relation to geopolitical factors shaping the world. Strategic intelligence help decision makers foresee and anticipate future events. By looking at 1514 strategic intelligence forecasts deducted from intelligence reports, they investigated the precision of these forecast. Results showed that forecasts within strategic intelligence were relatively precise in their predictions (Mandel & Barnes, 2014). They showed that forecasters displayed good discrimination, meaning that they had a good understanding of separating occurrences from non-occurrences'. Though, senior analysts were better at discriminating than junior analysts (Mandel & Barnes, 2014). This could be explained by the fact, that studies of judgmental tasks have showed that in order to judge a task, sufficient knowledge have to be present, meaning that judgemental tasks depend on the knowledge available to the specific individual (Nelson, et al., 1995). This study explored the accuracy of experts with specific field knowledge: However, other studies have investigated whether non-experts are more competent in developing forecasts than experts. Here it was found that non-experts are better at predicting future events than experts (Anderson et. al, 2005; Önkal & Muradoglu, 1994). In addition, a two-decade study of political judgements was conducted. Here it was shown that non-experts are at least or even better than experts in predicting future events (Tetlock, 2005).

3.4.1 Calibration:

Experts are shown to be better calibrated than non-experts in a number of cases including, outcomes of sports games (Andersson et al., 2005), predicting earnings (Whitecotton, 1996) and exchange rates (Önkal et al., 2003). In Mandel & Barnes' (2014) study of strategic intelligence the experts showed good calibration, meaning the degree forecasters subjective probabilities of their forecast is aligned with "*the observed relative frequency of event occurrences*" (Mandel and Barnes, 2014 p. 2).

"*An individual is well calibrated if, over the long run, for all propositions assigned a given probability, the proportion that is true is equal to the probability assigned"* (Lichtenstein et al., 1977, p. 552). Findings show that expertise is of importance and that practical analytical skills can be improved by experience (Mandel & Barnes, 2014). Other studies have also supported the fact, that experience gives experts the advantage of being better calibrated. Weather forecasters have been showed to be well calibrated when having to foresee rain (Murphy & Winkler, 1983), hockey players showed a good skill of calibration when having to foresee future matches' (Vertinsky, Kanetkar, Vertinsky, & Wilson, 1986)).

In Gideon Keren's (1987) study of bridge players, the results showed that expert players with more experience than those of amateur players, were better calibrated. As expert players expressed better calibration than amateur players, expert players were able to forecast the likelihood of a contract (a part of the bridge game) with more precision than amateur players. Amateur players portrayed a higher rate of overconfidence when having assigned high confidence to the ratings and vice versa. Additional, it was shown that amateurs showed a "optimism" bias, which experts were able to extract and therefore were more objective in their judgments (Keren, 1987).

Furthermore, the study looked into if items were related or not, meaning that the assessment of the probability assigned to an item can be related to experience within the field of the item (Keren, 1987). Thus, an experienced bridge player has played thousands of different hands that are independent of each other, these hands are still related to the same item/field, hence the same mental processes are used due to the similarity. Keren (1987) argues that good calibration constitutes of two sub-processes, where one is the process of semantic inference, where a person uses a mental model to generate non-numerical feelings of certainty from the person's own knowledge base. The second process is where the feelings of the first subprocess is transformed into numerical probabilities. Hence, it is argued that good calibration can be taught if these subprocesses are stimulated when the items are related. Thus, according to the findings, feedback and practice can teach better calibration (Keren, 1987; Tetlock, 2015). Conversely, to Keren's study (1987), Lichtenstein & Fischhoff (1977) show that psychologists are not well calibrated. Keren (1987) argues that this is due to lack of the second subprocess even though they have many related cases, they are not used to turn patients into probabilities.

3.4.2 Overconfidence:

Even though experts are shown to be better calibrated than non-experts, they are not necessarily well calibrated. A study of sports expert's predictions of the Soccer World Cup in 2002 showed a lack of calibration and a strong level of overconfidence (Andersson, Edman & Ekman, 2005). This was also seen in studies of a Russian manager's forecast of future economic (Aukutsionek & Belianin, 2001), and in a study of professional's prediction of future recession (Braun & Yaniv, 1992). Overconfidence tend to be high among economic predictions, as "*the combination of overconfidence and optimism is*

a potent brew, which causes people to overestimate their knowledge, underestimate risks, and exaggerate their ability to control events" (Kahneman & Riepe, 1988, p. 54).
One reason might be that information effects could play a misleading role for experts, as Anderson et al. (2005) showed that an increase in information is correlated with an increase in overconfidence, however, the accuracy of the predictions did not improve. Moreover, Goodman-Delahunty (2010) made a study of 481 lawyers in the United States in order to test their calibration level. The sample of lawyers were asked to set minimum goals of what they were to achieve on cases that were set for trial. Afterwards, they were asked to rate their confidence level of the likelihood of this to be achieved. Results showed that lawyers tended to be overconfident and that experience didn't change this fact. Furthermore, most research show that people tend to be inadequately calibrated (Lichtenstein, Firchhoff & Phillips, 1982). Fischhoff, Slovic, & Lichtenstein, 1977) showed that people in general are likely to be overconfident. They conducted five experiments, where the participants were asked questions of different levels of difficulty. After having answered the questions participants were asked to rate their own confidence level. All five experiments showed "*people tend to be wrong too often when they are certain they are right"* (Fischhoff, Slovic and Lichtenstein, 1977 – p. 561).

3.4.3 How to improve VC's decision-making:

The common assumption is that venture capitalists act as experts within the industry and therefore, most studies have looked into how venture capitalists evaluate start-ups. Thus, few studies have been concerned with how venture capitalists can improve their decision making or if any alternatives could increase the judgment process of assessing and investing in the most profitable startups. Though, within the literature of forecasting more emphasize have been placed on this subject in general. There is no clear generalization to be made, on whether statistical models outperform judgmental forecasts over a time series when the individuals have no knowledge of the field they investigate, as studies over time have showed different results. Hogarth & Makridakis (1981) concluded that "quantitative models outperform judgmental forecasts" (Hogarth & Makridakis, 1981p. 126). Though, in early-stage start-up investments, there is less amount of data to be quantified which leads to the usage of statistical models being limited (Appendix 7.1). In addition, Armstrong (1983) concluded that analysts who possessed domain knowledge were able to create far more precise forecasts than statistical models. Two reasons can explain this. First, statistical models become less and less accurate as time passes, hence, the knowledge of experts seem to be less time dependent and more up to date. Secondly, their information seems to be more comprehensive to the one of statistical models, as it also *"represents the un-modelled component"* (Lawrence et al., 2006 – pp.501). However, indications showed that a combination of statistical models and judgmental tasks of analysts could create an optimum, (Lawrence et. Al., 1986) as the combination could "*reduce forecast error*" (Lawrence et al., 1985 – p. 106).

Another way of improving VC's decisions is through the usage of feedback. Feedback is shown to have a positive effect on forecasters, as it can help them increase their performance (Benson & Önkal, 1992). Feedback within forecasting literature is known in many shapes. One type of feedback is known as outcome feedback - that simply refers to getting to know the result of the process. However, this type of feedback is shown to be the least efficient type of feedback (Lawrence et al., 2006). Whereas, other types of feedback, such as performance feedback is shown to help improve calibration of probability forecasts (Stone & Opel, 2000) as it more comprehensively explores the process of forecasters. Decomposition is also an approach to improve forecasting. Decomposition divides the holistic forecasting into a time series of smaller events. Few studies have investigated the effect of decomposition; however, it is shown to be better than holistic forecasting in some situations (Edmundson, 1990). On the other hand, Goodwin & Wright (1993) argued that decomposition can actually decrease the effect of predictions, when tasks are psychological complex, as it is the case to invest in early-stage start-ups.

Within venture capital, bootstrapping models are proved to have a positive effect on forecasts (Dawes, 1974). Bootstrapping models is one of the few approaches that have been investigated in relation to venture capital in order to improve decision making. "*Bootstrapping models attempts to capture the decision criteria used by the VC's in past assessments, and the relative weights placed of those decision criteria* (Shepherd and Zacharakis, 2002). Bootstrapping model's benefits decision as it express consistency and reduces information overload for the decision maker (Shepherd and Zacharakis, 2002). It improves the venture capitalists own understanding of their process and increases the fit between the decision-making processes of VC's and their systems of objectives and values. Another study found that the use of a bootstrapping model led to better predictions than judges own predictions (Dawes, 1975). Additionally, Zacharakis and Meyer (2000) showed that by using a bootstrap model and compare this to venture capitalists only one out of 53 venture capitalists matched the accuracy of the bootstrap model. As an extension to this model F. Csaszar et al. (2006) created a model that matched strategic and cognitive criteria. Hence, in the presence of limited information of certain specific questions of know-how or strategic perspective, this matter is investigated from a cognitive angle, providing the users with a way to reflect on their own answers. However, as bootstrapping models benchmark on prior decision criteria, which in the industry of VC's have not yet shown to be accurate, as venture capital are suffering from low success rates of investing in start-ups. Hence, evaluation criteria are argued not to be an exact measure of success.

Recent trends and studies have looked into the effects of machine learning and artificial intelligence (AI) within venture capital. As machine learning and AI have become better over time, methods that implement this, have received increased attention within venture capital. One of the first AI experiments that were conducted on AI in venture capital is from 2009. Here, journalists Ira Sager of Businessweek Magazine faced an AI company with the challenge of having to find the best 50 early-stage companies to invest in (Reddy, u.d.). The computer was given the opportunity to invest in all available start-ups in search for venture capital at the specific point in time. If the artificial VC was real and did in fact invest in the chosen companies, it would have been the second best performing VC until today (Reddy, u.d.). Moreover, in 2014, a Hong Kong based venture capital called Deep Knowledge Ventures gave an AI system a seat at its board (Quarterly, u.d.). Meaning that the AI system have the same saying and vote as the rest of the board members in Deep Knowledge Ventures. In addition, the tendency within venture capital seem to clearly point towards a more data-driven strategy to support human decision making. Veronica Wu who is managing partner at the data-driven venture capital Hone Capital says, that *"In the venture-capital world, success has historically been driven by a relatively small group of individuals who have access to the best deals. However, we're betting on a paradigm shift in venture capital where new platforms provide greater access to deal flow, and investment decision making is driven by integrating human insight with machine-learning-based models."* (outsideinsight, u.d.).

A cooperation between data-driven strategies and human decision insights are essential, as AI are still not able to feel or process feelings (Polanyi, 1966) (Hurree, u.d.). Petersen (2004) distinguishes between two different kinds of information, hard- and soft information. Soft information is more reliable on relationships (Petersen, 2004) e.g. chemistry between investor and the entrepreneurial team. Numerical information is thought to be hard information e.g. traction rates of a start-up. Hard information is not up for interpretation and only dependent on quantified data. Whereas soft information often is related to ideas, future plans and can often be interpreted (Petersen, 2004). Soft information is dependent on the collector and his interpretation of the obtained information. The purpose of soft information is not always clear and it can be hard to divide the relevant parts of soft information from the irrelevant parts. The meaning often does not reveal itself until the information is connected to a certain situation or looked back at in retrospective. The hard information is what AI's can interpret from as of today, hence the need for human capabilities to supplement AI is still essential. In alignment with Petersen's (2004) theory of hard- and soft information the literature of intuitive judgments has much in common with the idea of soft information. In general, intuition is knowledge we have that can't be explained. Polanyi (1966) presents the concept of tacit knowledge, by which he states that we "*know more than we can tell*" (Polanyi, 1966- p. 4). Tacit knowledge is connected to what we know, but it is not necessarily expressible. It builds on the know-how that people builds up over years and can thereby be said to be unique to each individual (Shirley & Langan-Fox, 1996). Tacit knowledge is closely connected to intuition, as intuitive decisions builds on the tacit knowledge individuals have stored (Hallin et al, 2009). Miller & Ireland (2005) introduces intuition as *holistic hunch* and *automated expertise* as two ways of categorizing intuition. Intuition as a holistic hunch is based on subconscious information gained from diverse and individualistic events and situations over time and is related to a feeling of knowing without knowing the explanation (Miller & Ireland, 2005). Whereas, automated expertise builds directly on relations to earlier subconscious situations and is a result of these learnings (Miller & Ireland, 2005).

Hence, as AI are not capable of interpreting feelings and other soft information, human capabilities is essential, especially in fields with high risk and uncertainty, which is the case of investing in early-stage start-ups. Generally speaking, early-stage investments are highly dependent on intuition and human judgements, whereas later stage investments can be supplemented with statistics, or as the theory states; Hard information. This study takes its focus on early-stage investments and therefore the part that is mainly involved with human judgments.

Chapter 4 - Hypothesis Development

The following section introduces the six hypotheses that are to be investigated. Each hypothesis is supported by relevant theories and previous studies. With these theories and studies in mind, the researchers will present their argument of the hypothesis development. The investigated hypothesis can roughly be separated into three categories. Hypothesis 1 and 2 test the total crowd and the experts vs. non-expert's ability to predict the success of the start-ups. Hypothesis 3 and 4 seeks to test the performance of small crowds of experts. The last two hypothesis seeks to test whether the crowd's express overconfidence in their predictions or if they are well calibrated.

Is combining knowledge the best?

Venture Capitals are able to invest in a large and diverse range of start-ups. Which of them that provides the best return of investment is, for even the best experts, difficult to predict. Venture capitalists are considered experts within the field of venture capital and argued to have prior knowledge of spotting the best investments (Zacharakis & Sherpherd, 2001). Early-stage investments, which is the focus of this study, is highly related to risk and uncertainty. In this stage, limited hard information is present and therefore it is a reliable strategy to depend on human intuition. However, the cognitive ability of humans can be restricted which leads to judgements being polluted by bounded rationality (Simon, 1987). Research has shown that collectives, in general, are better at solving certain task than individuals. In fact, Shaw (1932) found in his problem-solving test, that 53% of groups got it right, whereas only 7.9% of the individuals solved it correctly. Still, group work also has its limitations, as members of the groups might be biased by each other's opinion, which can lead to groupthink and biased decisions (Janis, 1972). Larrick and Soll (2006) found that increasing the size of a crowd in a prediction task will lead to an increased accuracy. In addition, Surowiecki (2004) argues that if crowds are diverse and independent of each other, they will be able to generate wise decisions. Galton (1906) did, in his experiment of a weight guessing task of an ox, discover something remarkable. The crowd which consisted of both experts with prior knowledge and non-experts, who found it fun to participate and come up with a guess, were in combination able to provide the most accurate answer. By aggregating human judgements, Galton (1906) figured that the misjudgements and prediction errors cancelled each other out. The phenomenon The Wisdom of Crowds has also shown to be generating the most accurate answer in other settings such as ranking weights or picking the order of horse race winners (Lee , 2012; Grazing, 2016). Hence, it is argued that an independent, large and diverse crowd consisting of experts within venture capital and non-experts with no experience within venture capital are likely to come up with an extremely accurate prediction. Thus, the following thesis states that:

H1: The total crowds' prediction is more accurate than the expert crowd's prediction

Does Diversity beat Ability?

When the effect of The Wisdom of Crowds is deployed under the right circumstances, a non-expert crowd consisting of a large-, diverse, and independent sample is proven to predict future events more accurately than few experts (Surowiecki, 2004). Large crowds with no specific field knowledge are able to come up with wise and precise forecasts (Surowiecki, 2004). In a two-decade study that investigated political judgments non-experts were shown to create more accurate forecasts than experts (Tetlock, 2005). In addition, studies of outcomes at the World Cup in soccer in 2002 (Anderson et. Al., 2005) and of stock prices (Önkal and Muradoglu, 1994) have both found non-expert's to be more accurate in their predictions than experts. Additionally, experts might suffer from information overload as they hold more information to shuffle between which can lead to biased decisions (Simon, 1987). Biases in general can act as an obstacle for experts through e.g. a tendency to generalize based on a to small sample, routed thinking etc. This criticizes the increased demand for expert knowledge and the tendency in many industries to 'Chase the expert' (Larrick & Soll, 2006).

Yet, other studies have proven that the knowledge and expertise of experts is in fact a reliable source for prediction (Mandel & Barnes, 2014). In the study of Gideon Keren (1987), it was shown how experienced bridge players were more precise in their prediction of future contracts. However, as Bridge is considered to contain less uncertainty and risk than the study of political judgement (Tetlock, 2005) and of stock prices (Önkal and Muradoglu, 1994), the researchers believe these studies to be more similar to the task at hand.

Therefore, the second hypothesis can be stated;

H2: The crowd of non-experts are more accurate than the expert crowd in their predictions.

Who are the better forecasters? Copenhagen vs. London

"The London startup scene is the biggest in Europe and it's growing faster than any ecosystem in the US. It's the most diversified startup ecosystem in the world. It has an exceptional market reach and you can meet the most affluent clients and investors there." (Ciszewski, 2016)

Evidence have shown that the London Ecosystem is way ahead of Copenhagen in terms of performance. However, Copenhagen has a large potential for future growth as they rank among the best countries for producing start-ups. Apart from the Ecosystems, which is to some extent also dependent on the size and location of the country, studies have shown that VC performance and success is highly related to cultural difference (Nahata, 2014). The two main differences in terms of Ecosystems is that Denmark is struggling by producing scale-ups compares to London (MandagMorgen, 2015). As illustrated in Figure 4, cultural difference is also highly present. Figure 4 illustrates the 6 dimensions of Hofstede when applied to Denmark and United Kingdom. Important to note is that these dimensions are regarding countries, whereas Ecosystems are focused on the cities, London and Copenhagen. The only element where Denmark is marginally in front of London is indulgence, meaning that people are generally willing to fulfil their needs with regard to enjoying life. The low score in masculinity – in other words, a feminine country, is highly related to a society where the success is determined by quality in life and the motivation to stand out from the crowd and be the best is not the ultimate goal (Hofstede, 2011). Opposite to London, who has a masculine society, the desires are focused on performance and competition. This element is particularly related to the performance of the countries Ecosystems.

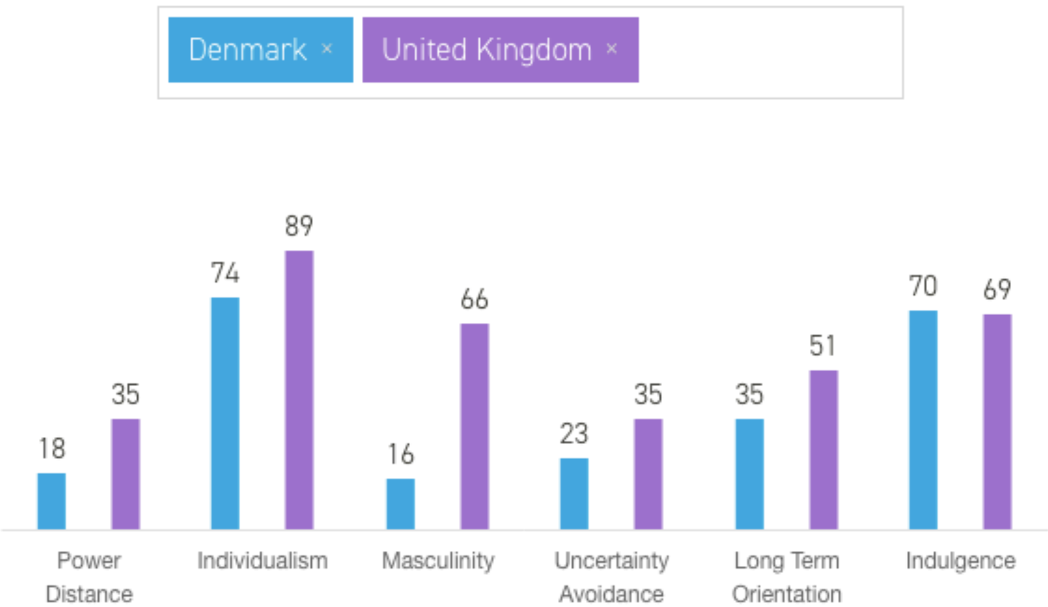


Figure 4 - Source: (hofstede-insights, 2018)

As VC's and investors are a key driver for fostering growth of the Ecosystem, it argued to be of interest to investigate whether experts in London are better than Danish experts at spotting the highest potential. London has proved its capabilities of supporting startups, creating a rich and entrepreneur friendly environment and managing to produce billion-dollar scaleup companies such as Powa, Shazam and ARM Holdings (StartupGenome, 2017). As Denmark is among the leading countries of producing startups (MandagMorgen, 2015), it is argued to be a key motivator for experts in Danish VC's, to increase the success rate of startups and to maintain growth of start-ups inside the country. The next hypothesis to be investigated is therefore whether foreign VC's are better at predicting the success of a startup than Danish experts.

H3: The prediction by Foreign experts is more accurate than the prediction by Danish experts.

Does experience make you a better forecaster?

Mandel and Barnes (2014) investigated the accuracy of forecasting within strategic intelligence. By looking at 1514 forecasts deducted from intelligence reports, the results showed that the predictions of strategic intelligence were relatively accurate. This indicates that in some industries, the importance of expertise and experience is key when forecasting future events. Additionally, did Franke et al. (2008) show that evaluation within venture capital correlated with experience. Thus, VC's that have more experience than their colleagues tend to add value to characteristics of a start-up differently than their less experienced colleagues (Franke, Gruber, Harhoff, & Henkel, 2008). However, Shepherd et al. (2002) showed that experience is only beneficial to a certain maximum level. Too much experience can lead to heuristic shortcuts and other biases such as routed ways of thinking, overconfidence in past decisions and approaches (Simmons et al., 2003). Though, experience was showed to be beneficial to a certain level, as information overload can occur when experts seem to have insufficient experience (Shepherd et al, 2002). This is consistent with the fact, that in order to judge a task a certain level of knowledge have to be present to the forecaster (Nelson et al., 1995). In other words, a prediction is dependent on the predictor and his level of knowledge. Findings show that individuals are able to develop their accuracy ability through exposure to work tasks of related items (Keren, 1987). Hence, working with tasks that are related to the same judgement process can increase accuracy. As venture capitalists are experts within their field, and work in a constantly changing environment, the researchers believe that they are less likely to be influenced by heuristic shortcuts, as their field is constantly changing, which is a factor they have to keep in mind and adapt to. Additionally, it is argued that judging startups are related tasks, even though the start-ups venture capitalist's asses can be completely different. Hence, it is believed that experience within venture capital increases the prediction accuracy.

H4: Experienced venture capitalists are more accurate in their prediction than less experienced venture capitalists.

Are Experts better calibrated than non-experts?

Calibration "*refers to the association between judgments of objective and subjective probabilities of the occurrence of an event"* (Weber & Brewer, 2004, p. 157). Good calibration is generated from two sub-processes of generating feelings of certainty towards an item based on one's knowledge base and then finally converting this feeling into a numerical probability (Keren, 1987). Studies of sports games (Andersson et al., 2005), earnings (Whitecotton, 1996), exchange rates (Önkal et al., 2003), bridge players ability to foresee contracts (Keren, 1987), and weather forecasters ability to predict rain have all shown that experts are better calibrated than non-experts. Experience gives experts the advantage of connecting patterns of related items and thereby developing their ability (Keren, 1987). Additionally, experts are shown to be better relative to non-experts to subtract optimism biases from their predictions (Keren, 1987). However, other studies have shown that experts are not likely to be well calibrated, as Lichtenstein et al., (1982) showed in his study of physicians. Nonetheless, it is believed that venture capitalists are more likely to generate probability metrics for their cases than physicians are of their patients. Physicians have to treat all their patients to their best ability, not predict whether they are to become well. Whereas, predicting probabilities is an essential part of a venture capitalists' job. Venture capitalists are presented to plenty of start-ups from which they have to invest in the best ones, much like the studies of bridge players (Keren, 1987) and sports games (Andersson et al., 2005). Additionally, venture capitalists should become better over time to generate probability outcomes from their knowledge base, as it will be increased over time, as startups are argued to be related items.

H5: Experts are better calibrated than non-experts

Are predictions victim of confidence bias?

Studies of sports expert's forecasts at the World Cup (Andersson, Edman & Ekman, 2005), Russian manager's forecast of future economic (Aukutsionek & Belianin, 2001), and professional's prediction of future recession (Braun & Yaniv, 1929) have all showed that experts tend to be overconfident in their predictions. Overconfidence is relatively often seen within economic predictions as many predictors hope for an optimistic outcome, causing an overestimation of one's own knowledge, resulting in extreme overconfidence of the specific predictions (Kahneman & Riepe, 1988). Information effects have been shown to play a distractive role for experts, as they tend to be more overconfident consistent with the more information they possess, however, this is not reflected in their predictions (Andersson et al., 2005). Nonetheless, it is not only experts that tend to be overconfident. Slovic et al. (1977) showed that individuals in general are overconfident as "*people tend to be wrong too often when they are certain they are right"* (Slovic et al., 1977 – pp. 561). Nevertheless, findings have showed that there exists a positive relation between overconfidence and experience as more experienced individuals tend to be more overconfident than less experienced individuals (Fazio & Zanna, 1978).

Hence, the researcher expects the results of this study to show overconfidence in predictions among both experts and non-experts. However, it is expected that experts will place a higher degree of subjective certainty to their objective predictions than those of novices.

H6: Both experts and non-experts are overconfident in their predictions, however experts are relatively more overconfident than non-experts.

Chapter 5 - Methodology

The following section is based on the work of Alan Bryman and Emma Bell's, Business research methods (Bryman & Bell, 2003). The section will give the reader insights of the methodology of the thesis.

The field studied and the problem statement sets the methodological approach of this study. The objective of this study is to investigate whether the theory of wisdom of crowds, that is a part of collective intelligence is an effective method to increase predictions of which start-ups to invest in within venture capital. The researchers aim to understand whether experts have a role to play within venture capital or if crowds act more sufficiently in predicting than experts. Additionally, the researcher examines whether different constellation of crowds is more effective than others, hence, if experts with specific characteristics are more accurate in their predictions than others. Finally, the study explores whether experts and non-experts are well calibrated when having to create a forecast of a hypothetical investment.

The researchers hold an epistemological position of positivism. Thus, believe that the methods of the natural sciences can be applied to study the social reality (Bryman & Bell, 2003). This is an effective position to embrace in this study as science generally is deterministic meaning that X, causes Y (Bryman & Bell, 2003) - and the role of the researchers is to uncover the relationship of cause and effect. Consequently, this study investigates whether large crowds have a certain effect on predictability of start-ups' success and if rules can be developed and generalized from these observations. Additionally, as the social reality is external humans are seen as natural objects whose behaviour are explained by external forces they have been exposed to (Bryman & Bell, 2003). Hence, the biases incorporated in individual actors that Wisdom of Crowds argues to remove, occurs due to external forces that acts on humans. As empiricism is part of science, the study only deals with what can be seen and measured. Hence, the study exclusively focuses on facts in findings and exclude interpretations of the data. As the findings of the research study can be measured, these can act as facts and therefore, the researchers can set themselves in a objective position when analyzing the data.

The researchers hold an ontological position of objectivism. Thus, the researchers believe that a social reality, that is external to social actors, exists (Bryman & Bell, 2003). The objective of determining whether collective intelligence can be used to increase the process of picking successful start-ups within venture capital can either be confirmed or rejected, but not constructed by the social actors, which also correlates with theory of collective intelligence that the average of the respondents' answers is the best possible answer. Hence, the study holds an approach of objectivism "*that implies that social phenomena confront us as external facts that are beyond our reach of influence"* (Bryman & Bell, 2003, p. 19). Thus, the theory of collective intelligence cannot be influenced by the individual actors.

As the objective of this study is to investigate whether collective intelligence can optimize the process of investing in start-ups within venture capital, the researchers acknowledge that collective acts as established knowledge and the premises of a deductive approach (Bryman & Bell, 2003). A deductive approach is well suited for this study as we hereby move from generalizations of forecasting to testing the specific case through data collection within venture capital. Most literature has focused on judgmental forecasting, however, none have specified it and tested it within the settings of venture capital. The study is deductive as it had its point of origin in a thoroughly review of the existing literature and theories on collective intelligence that have been studied in many different relations and within different fields. From this, hypothesis' have been deduced on how the theory of Collective Intelligence and Wisdom of Crowds was likely to act within the field of venture capital. Thus, a deductive reasoning process was conducted as theory have guided the research and not the opposite way around. Though, the final process of the research study transformed into an inductive process, as this was where the findings and their implication for the existing literature were exploited.

5.1 Research design

As the study seeks to investigate whether a large diverse crowd is a more accurate predictor than a small group of experts, if experts from London are better predictors than experts from Denmark, whether experienced experts are more accurate in their predictions than less experienced experts and if experts are well calibrated compared to non-experts, a comparative research design is chosen. This enables the researchers to compare the data of the different groups to each other as "*this design entails the study using more or less identical methods of two or more contrasting cases"* (Bryman & Bell, 2003 p. 56). A comparative design "*embodies the logic of comparison in that it implies that we can understand social phenomena better when they are compared in relation to two or more meaningfully contrasting cases"* (Bryman & Bell, 2003, p.56). The comparative design lies close to the crosssectional design, as it is measuring on different factors at one point in time (Bryman & Bell, 2003). Though, the comparative design is chosen, due to its applicability of comparing cases. Hence, one part of the comparative design will solely focus on cross-national comparison of the two groups, as London is known for having a better ecosystem and significant differences in than Copenhagen. Another part will focus on comparing the large crowd of non-experts to the small crowd of experts to measure the effect of predictability - thirdly a part will focus on comparing the effect of experience within experts and finally. The last part will focus on whether calibration within capital is correlated to having prior expert knowledge.

Strategy for literature review

The foundation for this study was the extensive literature review. Hence, the strategy for collecting the right literature is essential to the outcome of the study. The objective of a literature review is to establish what knowledge, thoughts and ideas that exists within a specific area of topic. The process of selecting the right literature is essential to do a valid and trustworthy analysis (Latsis, 1981). In order to select the right literature, the researchers were aware of picking acknowledged and peer-reviewed articles. To collect the literature, databases such as CBS's Libsearch and Google Scholar were intensively searched through. Additionally, Google search was used to get the most recent articles of trends within venture capital, as many of these trends are yet to be studied by scholars. However, when articles were used from alternative websites, these were only used if the researchers had confirmed the knowledge retrieved on different webpages and from different perspectives.

The study focused on three areas of literature:

1) The first area explored the literature of Collective Intelligence. The researchers have had a course that among other fields were focused on Collective Intelligence, hence the curriculum from this course was explored.

2) The second area that was explored was the literature of Venture Capital. Here the main focus was to explore what criteria venture capitalist's asses' companies by and exploring trends and studies that have tried to optimize the process of venture capital, which is the objective of this study. As most studies of venture capital focused on evaluation criteria when assessing which start-ups to invest in, few studies were concerned with how to optimize the process of venture capital, hence, findings of some of the newer trends within venture capital did not come from scholars, but from people with practical experience within venture capital. However, to make sure the data was reliable only articles

by respected individuals within venture capital was used. Additionally, the researchers collected primary data from key figures within well renowned Danish venture capitals, in order to validate the findings from the literature. Thus, the study used both primary and secondary data, which will be more thoroughly examined in the following section.
3) Lastly, literature within judgmental forecasting was collected. Primarily, this field was investigated as the task of the thesis is to predict a future event. In addition, literature of overconfidence and ways to improve decision making is explored to understand the underlying factors that might affect investment decisions.

5.2 Research strategy

The study has chosen a quantitative research strategy (Bryman & Bell, 2003), as the purpose is to measure and compare findings and ultimately generalize these to the population through a test of significance. With this in mind, the study still used both qualitative and quantitative data gathering. Yet, the qualitative data is only used for the purpose of building up the cases, which is eventually measured upon. Hence, the results stem solely from the quantitative data. However, the method of using both quantitative and qualitative data is known as **triangulation** (Bryman & Bell, 2003). Triangulation was used to take as many perspectives as possible into consideration. Thus, triangulation had the effect of validating findings from different angles, ensuring the quality of the findings.

5.3 Qualitative data

The purpose of the qualitative data was support the development of the investigated hypotheses. The qualitative data were conducted as semi-structured interviews and were used to combine practical relevant factors with the findings from the existing literature. The interviews were semi-structured and the interview guide (Appendix 4) were developed in accordance to existing theory on judgmental forecasting and collective intelligence. Semi-structured interviews were used as these facilitate an open framework for two-way communication within the scope of the field studied (Bryman & Bell, 2003). Thus, the interview guide consisted of general questions based on findings in the literature of venture capital, but due to the selection of semi-structured interviews the interviewers were able to ask follow-up questions when interesting responses were identified (Bruton & Ahlstrom, 2003). Additionally, semi-structured interviews conducted face-to-face increases the quality and honesty of the interviews (Bruton & Ahlstrom, 2003). The qualitative interviews consisted of three interviews of approximately 30 minutes each with investment managers of large Danish venture capital funds. These were conducted in order to understand which criteria venture capitalists evaluate startups by when assessing whether to invest in them. Though, criteria were already discovered from the literature review, the interviews were conducted to make sure the findings from the literature corresponded with real practice. Additionally, as no recent studies have investigated whether Collective Intelligence can be used to forecast the investments of venture capitalists, the interviews were used to establish if procedures of Collective Intelligence were already applied in VC's. Finally, the interviews were used in a preliminary sense, in order to construct the most appropriate scope of the study.

The interview guide to qualitative interviews

The interview guide for the semi-structured interviews was based on the chosen fields of literature (Appendix 4). The interview reflected the knowledge obtained from the prior literature review and was overall divided into five subcategories. The first part regarded criteria that venture capitalists evaluate start-ups by. The second part was concerned with the decision-making process of investing and if the process reflected main characteristics of James Surowiecki's (2004) theory on Wisdom of Crowds. The third part was concerned with the relation of- and distribution of experience within venture capital in alignment with studies of etc. Mandel & Barnes (2014). The fourth part was concerned with how venture capitalists relate to their own process and how they seek to become better at forecasting in regard to Tetlock's (2015) study of Superforecaster's. The final part explored to term success within venture capital. How success is perceived and dealt with.

Sampling for interviews

In order to generate interviews, the researchers used simple random probability sampling. The researchers contacted, to their best knowledge, all venture capitals located in Copenhagen. Three venture capitals responded, ByFounders, Vækstfonden and PreSeed Ventures, which resulted in three interviews with investments managers from different management levels.

Findings of the qualitative interviews

Analysis of the interviews was done right after the interviews as Bryman & Bell (2003) recommends. The authors choose to pick only relevant citations from the three interviews which is in accordance with the findings of the literature. Moreover, the authors categorized the analysis from the interview into the evaluation criteria of VC's, the usage of Collective Intelligence and general judgmental initiatives in VC's. The full interviews are attached as three mp3 files and a brief summary of each interview can be found in Appendix 7.

5.4 Quantitative data

Quantitative data were the main focus of this study and were used to measure and compare the findings and to test for significance. The quantitative data were gathered in order to measure the effect of Wisdom of Crowds when applied to predicting the success of start-ups. The data were gathered through a questionnaire, that acted as the crowd prediction tournament. The questionnaires were online and made through Google Forms. The questionnaire consisted of three parts. The first part was made to generate data about the respondents in order to ensure diversity in the crowd and measure different demographic variables in correlation to each other.

The second part consisted of four historical startup cases that the respondents were asked to thoroughly read through. All four cases were presented as a two-page PDF, that provided information of the product and team of the startup. To ensure that none of the respondents were able to identify the startups all names of the company and team members were erased from the two-pager. Additionally, respondents were asked not to respond to the case, if they were able to identify the company behind it anyway and had prior knowledge of them as an investment case. All four startups had received funding from Almi Invest, one of Sweden's largest venture capitals, with different outcome of exit. One of the cases had a return of investment of 10x (which is the success criteria of venture capitals) and the others had respectively a return of investment of 8x, 3x and 0x. The criteria that the cases were based on were carefully selected on the foundation of a comprehensive literature review and through interviews with Danish venture capitalists. Information on the startups given to the researchers by Almi Invest. Additional information that was needed to build the cases were gathered from LinkedIn, company websites, information from the original founders and databases Pitchbook, Dealroom and Crunchbase (Databases of startups). The researchers tried to make the cases as similar as possible in sense of the information given to the respondents.

In the third part of the questionnaire respondents were asked to rate the cases through different methods. First, a 5-point Likert-scale were used to rate the four different startups on two measures of "quality of product" and "quality of team whereas the value 1 was the worst quality and 5 was the best possible quality. A Likert scale were used as Blohm et. Al. (2016) found that when investigating which rating scales are best for evaluating open idea evaluation, rating scale-based task raises ease of use for the participants and increases the decision quality of the participants. Hence, mechanism

accuracy of aggregating individual answers within Collective Intelligence is not just an indication of the aggregation mechanism, but also affected by the decision taken towards the user level. Thus, ease of use for the participants increases their decision quality as "*ease of use frees cognitive resources and allows users to make more accurate idea evaluation decisions"* (Blohm et Al., 2016, p45).
Secondly, the respondents were asked to distribute/invest 1.000 points in total in the four companies, in the way they believed to be the best investment as possible. They were told, that they could invest as much as 1.000 points in one case and 0 points in the rest or distribute the points in some or all of the companies. This was fully up to the respondents to decide. Before investing, the respondents were also told that the companies had given different rate on return of investment and that these rates were 10x, 8x, 3x and 0x.
Finally, the respondents were asked on a 10-point Likert scale to rate their own confidence level of their investment. Moreover, the respondents were asked to put a few words of why they chose to invest as they did. This was done to investigate whether respondents showed over- or under confidence in their investment/prediction.

Framework

In order to create consistency among the cases presented in the crowd prediction tournament, the researchers developed a framework of evaluation criteria that should be present in the four cases. The framework is developed from existing literature of evaluation criteria within venture capital and from the interviews conducted by the researchers with investments managers from large Danish VC's. The researchers have identified three general categories of evaluation criteria that can be developed – *team*, *product* and *market*. These categories are backed up by existing literature (Wells, 1974) and the interviews (Appendix 7; Appendix 7.1). The market is an extremely valuated evaluation criteria within venture capital, as the rule of thumb is that a start-up should have the potential of creating a return of investment of at least 10x (Lemkin, 2012). Hence, for the market to be interesting as a venture capital case, it potentially needs to be a billion-dollar market (Appendix 7). As all the start-ups presented in this study have received funding, the assumption is that the market potential is high and reliable as venture cases. Additionally, the researchers would not be able to conduct historical data on the four specific markets, thus, this criterion is left out of the framework. The team evaluation criteria are established in alignment with Franke et al. (2008), that have looked into how evaluation criteria of the team within venture capital are weighed. The product evaluation criteria are established in alignment with Song et al. (2008), Bruton & Rubanik (2001), Macmillan (1985) and the interviews. The framework is presented below:

Framework of evaluation criteria within venture capital

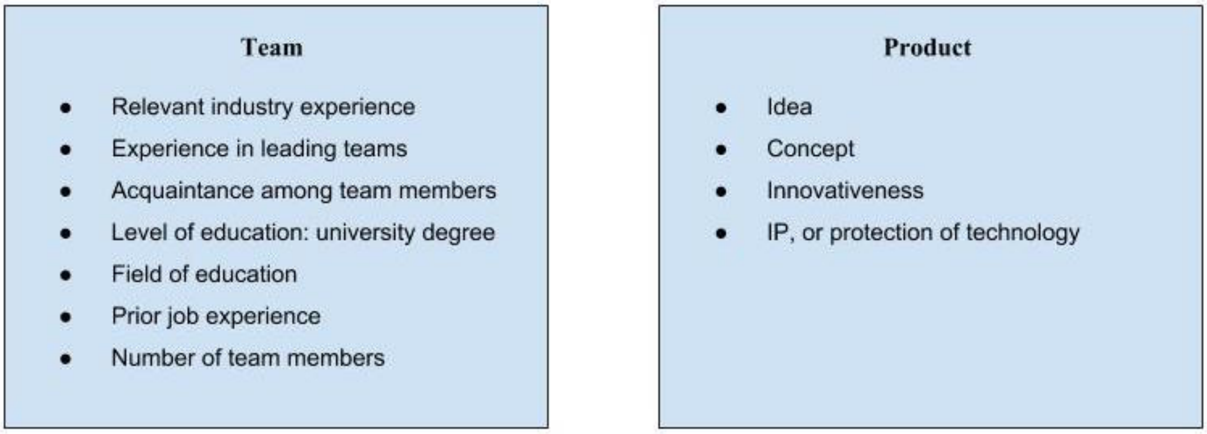


Figure 5 – *Source: Own work*

Sample of non-experts

In order to generate respondents of non-experts for the questionnaire the authors got three tutors from Copenhagen Business School to disperse the questionnaire on their specific study's Facebook page. We have no connection to the people that the questionnaire is disperse to through the tutors. Additionally, public Facebook pages for dispersing questionnaires was used. Additionally, the researchers placed themselves at "Strøget" in Copenhagen with free candy that people could take, if they promised to go back home and answer the questionnaire. Here people were told about the project investigated and the researchers just relied on people's good faith, when saying that they were to answer the questionnaire. No information specific information of the task at hand were given.

Sample of experts

In order to generate answers from Danish experts the researchers have reached out to, the researcher's best knowledge, all VC's based in Copenhagen. From this random sample, the authors achieved an expert crowd consisting of four experts from Danish VC's. In order to generate respondents of foreign experts, the researchers e-mailed approximately 30 VC's placed in London, which was drawn from a list of all VC's London. This resulted in three foreign experts who worked in a London VC. As email correspondence were conducted with the experts, these were given some additional information about the project, but not any knowledge that had to do with the questionnaire. Hence, the experts had the same knowledge of the specific cases as the non-experts.

To keep the crowd tournament interesting and attractive to the participants and furthermore make the respondents answer to their best ability, the researchers set a prize of 500 DKK / 70 EURO for the most accurate prediction, which is measured through the highest return of investment. Financial rewards are considered a relevant factor for motivating the crowd (Malone, 2009).

5.5 Quality of quantitative data

Many precautions and considerations have been made in order to ensure the quality of the data. In order to ensure replicability and generalizability if the study, the following section will investigate the reliability and validity of the quantitative data.

5.5.1 Reliability

Reliability is concerned with whether the measures of the study are consistent. Hence, the focus is on whether the same findings would occur by repeating the study. Reliability is constituted of three factors; stability, internal reliability and Inter-observer consistency (Bryman & Bell, 2003). In order to ensure that measures of this study are consistent, the researchers developed a framework to their best ability from existing literature within venture capital and from interviews with investment managers from VC's, to make sure that the factors presented in the cases reflected the criteria investment managers within venture capital asses' start-ups by. As this framework is developed from literature reflecting almost the entire history of VC's, the measure should be stable, as venture evaluation criteria have been consistent over time. Hence, if the study were to be conducted again, the measure of Collective Intelligence's impact on venture capital should be possible to be reconstructed.

The internal reliability reflects whether key indicators of the survey are consistent (Bryman & Bell, 2003). Internal reliability is measured by seeing if there is correlation between the measure of the total score at the Likert-scale and the corresponding investment of points in the four companies. Meaning that the way respondents rank the companies on the Likert correlates with how they rank the companies through their investment. However, it should be noted, that the respondents are not specifically asked to rank the companies, though it could be anticipated that the correlation isn't going to be perfect. Yet, the participants' assessments of the start-ups are likely to reflect a ranking anyway that should create a correlation. Through the results, it was found that, in terms of ranking, the Likert-scale rating significantly corresponded with the ranking of the investments. Hence, it is argued that there is internal reliability, as participants understood the internal linkage between the tasks. Thus, internal measures are consistent with each other. The inter-observer consistency is believed to be high as limited judgmental tasks have been involved in developing the research, as the study mostly follow guidelines from earlier conducted research studies and that the researchers have weighted the arguments for and against the different procedures, in order to come up with the best solution for this study. In general, the researchers have agreed on what they measured.

5.5.2 Validity:

The validity of this study is concerned with "*whether a measure of a concept really measures that concepts"* (Bryman & Bell, 2003, p.77). To make sure that the measures of the study at least reflected the concept of attention, also known as the face validity, the research team sought guidance through different sources. In order to secure the validity of measures reflecting the part of Collective Intelligence the researchers looked at other relevant studies that have measured similar effects and got inspiration from the ones most suitable for this research. Additionally, the researchers corresponded with their supervisor, who holds comprehensive knowledge within the field of Collective Intelligence. In order to ensure the validity of measures of venture capital, as earlier mentioned, a literature review was conducted. Additionally, the researchers conducted interviews with key figures within the industry of venture capital, to make sure that the survey would reflect the same criteria that venture capitalist usually evaluate start-ups by. Hence, the results are believed to reflect the concepts it measures.

5.5.3 Replicability:

In order to produce a high degree of replicability the researchers have tried to describe the procedures of this study in great detail. And believe that the already abovementioned of how data have been retrieved explains this extremely descriptively. Consequently, the results of this study should easily be replicable to other researchers, who wishes to investigate the subject at matter.

5.6 Quality of qualitative data:

In order to ensure the quality of the qualitative data the following section will account for the trustworthiness and authenticity in accordance with Lincoln & Guba (1985) and Guba & Lincoln (1994).

5.6.1 Trustworthiness

Trustworthiness in Guba & Lincoln's (1994) version consists of four subcategories;

1. Credibility that refers to how believable the findings are.
2. Transferability that refers if the findings are applicable in other contexts.
3. Dependability that refers to whether the findings are consistent over time.
4. Confirmability that refers to whether the researchers have succeeded in staying objective.

The researchers believe to their best knowledge that the trustworthiness of the findings from the interviews are high. The researchers found that a high consistency between the three interviews existed and the findings corresponded with the key findings from the existing literature, hence the credibility is believed to be high.

The interviews were with three investment managers from three different venture capitals in Denmark, who is argued to be a rather precise representative sample of the Danish venture capital environment. Thus, the transferability is argued to be high.

In order to have a high degree of dependability the researchers have sought to take an "auditing" approach (Bryman & Bell, 2003) and provided comprehensive descriptions of the different phases related to the interviews. More comprehensive records can be found in the Appendix 7. Finally, the researchers have to their best knowledge stayed objective in the interview setting and ensured a high degree of confirmability. This was an easy process for the researchers, as the interviews were used to build the cases the study is based on, therefore, the researchers had no interest in affecting the interviewee with subjective values and believes, as this could lead to a flawed structure of the cases presented in the study.

5.6.2 Authenticity

Authenticity consists of the five subcategories *fairness, ontological authenticity, educative authenticity, catalytic authenticity and tactical authenticity* (Guba & Lincoln, 1994). The interviews are believed to hold a high degree of fairness as they fairly represent the social setting of VC's in Denmark as three different investment managers from different venture capitals are the interviewees. In addition, the interviewees present different managerial levels differentiating from a CEO to investments managers with different experience levels. Thereby the researchers have tried to represent different perspectives of the venture capital environment. Ontological authenticity is present as members will gain a better knowledge and understanding of the VC milieu from the research study. The catalytic- and educative authenticity are also believed to be present, not necessarily to a high degree but as the findings of the study are not going to be confidential, these should be present to a certain degree. The researchers have been asked by the involved venture capitalists to share their results with them, hence tactical authenticity is also present to a certain degree.

5.7 Limitations

The research team are aware that the study holds some limitations. Firstly, the investment cases used for the questionnaire were historical, and are therefore being assessed in another time than when they got their initial investment from Almi Invest. This means that trends of innovation could have developed over this time period and therefore affected the predictions of this study. Additionally, information of the companies was retrieved from what was accessible to the researchers, which changed from case to case. Though, it is argued that this doesn't affect the reliability of the sample, as the population of the sample were all given the same information and the researchers build the cases from a framework in order to create consistency between the information given in the different cases. Yet, the information regarding the start-ups could've affected how the sample were to predict, as the changes in the given information could change the impression of the crowd. Additionally, the four cases are all B2B (Business-to-business), which could be argued to make them more difficult for the crowd to understand them. This could especially be favourable for the experts, as these are likely to easier relate to B2B cases than the non-expert crowd. Furthermore, the literature review and the interviews have showed that the team behind the start-up is of extreme importance when venture capitalists asses which companies to invest in. In this research study the team was presented on paper, hence, it is difficult to get a comprehensive feeling of the persons behind the companies.

Lastly to mention, there is a limitation to the study in terms of the task of investing 1.000 points.

In the general approach of developing the optimal investment portfolio, risk is a large part of the equation. However, the risk factor is not present in the task of investing in the four start-ups as the risk is considered the same. Still, many participants intuitively spread their investments in order to account for risk.

5.8 Data analysis

To test the hypothesis developed by the authors, an analysis of the data is required. In the existence of many different methods for explaining data, the authors have chosen methods that are in alignment with previous studies on Collective Intelligence and The Wisdom of Crowds (Larrick and Soll, 2006; Blohm et al., 2016; Surowiecki, 2004) The choices ranges from descriptive statistics to inferential statistical analysis. The following section seeks to explain chosen models and how the analysis for this project was done.

Before moving into the introduction of the selected models, two important notes for this study is:

1. The participants of the crowds are only asked to evaluate the cases once.
2. As the cases presented are all historical, the researchers know the "answer" on beforehand. The "answer" are the return of investments the start-ups have provided.
3. The participants of the crowds are independent of each other and therefore, there is no possibility for information sharing to affect the results or the possibility to modify answers based on comparisons to the other participants.

Choice of Models

Firstly, to analyze and develop a comprehensive understanding of the data from the sample the researchers present the descriptive statistics. The purpose of descriptive statistics is to provide the reader with a brief summary of the sample. Furthermore, it is a good way of visualizing the overall performance of the responses (Bendsen, 2018). The descriptive statistics will visualize the variables of the experiment and the proportion of these (Bendsen, 2018). The variables included in the research are gender, age, nationality, educational background, employment status, and whether the participants held experience within venture capital and by that, could be considered as an expert or non-expert.

Correlations

In addition to this, the authors will present a correlation matrix which will be presented in Appendix 2. A correlation matrix is a good measurement for the strength of the relationship between two or more variables in the data (Bendsen, 2018). Correlation ranges between -1 and +1, where a +1 is referred to as a perfect correlation and -1 is when two variables are imperfect correlated (CreativeResearchSystems, 2018). As a comparative study is conducted, correlations between the crowds and the quality rating and specific investments are interesting. Note that correlation only captures linear dependence in the data. Non-linear correlations are not captured by the simple Pearson correlation (CreativeResearchSystems, 2018). Additionally, a correlation matrix between the investment task and the rating scale task will be produced, as this will act as a way to secure internal reliability.

Regression analysis

To test for significance of the correlations, the authors will conduct a multiple regression analysis between e.g. the dependent variable, *Investment* and the two independent variables *Product* and *Team*. The multiple regression analysis describes the relationship between the independent variables and the dependent variable (Bendsen, 2018). Adding more independent variables can create multicollinearity, which is an expression of the independent variables' correlation (Graham, 2003). To overcome multicollinearity, the authors conduct both multiple regression and simple regression analysis to ensure the correlations are right.

The analysis will provide the researchers with a P-value that expresses the significance of the results. The P-value of this study has to be <0,05 in order to be considered significant. Moreover, the regression analysis offers a Multiple R, R-Squared and Adjusted R-square which are indicators of the strength between the data and the regression line provided. Multiple R is the correlation coefficient and expresses how strong the linear relationship between X and Y is. R-Squared is used in the simple regression analysis and explains how differences in one variable can be explained by the differences in the second variable (Bendsen, 2018). Adjusted R-square, however, displays the accuracy of the dependent variable when terms of the model are accounted for, hence, it only increases or decreases when the new variable have an actual effect on the predicted value (Bendsen, 2018). Thus, it is used in when multi-regressions are made.

Chin (1998) identified how to analyze certain levels of adjusted R-square. He considered 0.67, 0.33 and 0.19 to be respectively substantial, moderate and weak.

Aggregating

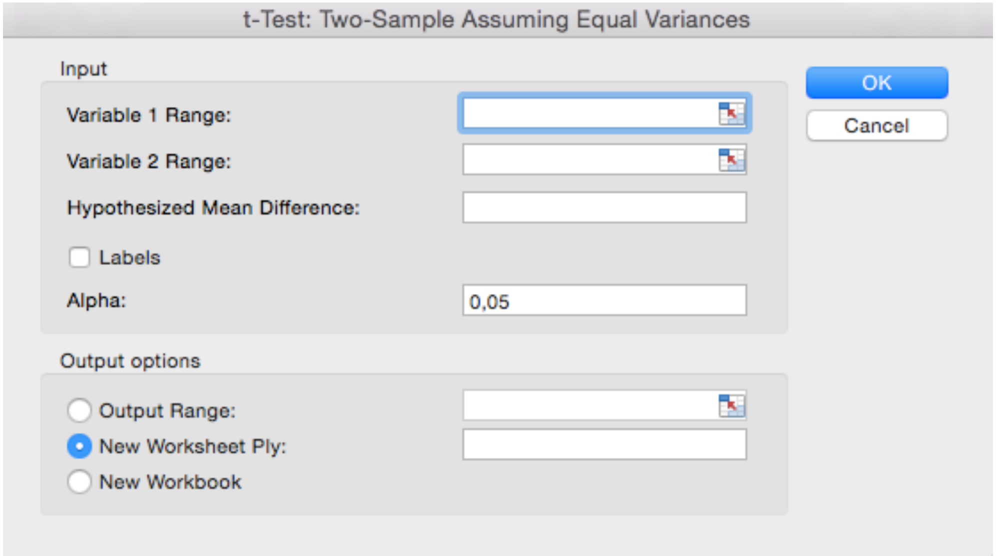
One of the most common ways of aggregating judgments from crowds is by using the *unweighted linear averaging* (Lyon & Pacuit, 2013). This model was used by Galton in his experiment of independent individuals guessing the weight of an Ox. Since then, the model has been considered a gold standard for aggregation by many researchers (Armstrong, 2001; Keuschnigg & Ganser, 2017). In addition, Armstrong (2001) has recommended this model as a good default model if there is no prior information about the crowd participants. If previous information about forecasting skills and expertise is present, the *weighted linear average* is a reliable alternative (Lyon & Pacuit, 2013). The drawbacks from averaging can occur under certain circumstances, e.g. large outliers might skew the mean. If the individual judgements are clustering around a central value *unweighted linear averaging* is a suitable approach. Moreover, if a Likert-scale to determine quality is present, averaging is a good method for getting an overall perceived value for an item (Malone, 2009; Blohm et al, 2016)

Statistical Hypothesis Testing

In order to secure robust conclusions, the researchers need to test the statistical hypotheses. Based on sample statistics, the researchers would like to make informed guesses about the true population based on the data from the extracted sample. The most basic underlying idea is to test whether the observed statistics is due to pure randomness.

Two different tests are carried out by the researchers. First, a test for difference in sample means and secondly, a test for differences in sample proportions. The first mentioned is a test of continuous output where the other one tests a categorical output. In other words, the test for differences in sample means is used to compare the return of investments for e.g. the experts versus the total crowd. The categorial test is used to test whether the crowd is better at picking the best investments compared to the experts. The 'best' is here defined to be true if the investor has invested most in Company 1 and/or Company 4. As these two start-ups are the best investments, a categorial test will be conducted on each start-up.

To carry out the test for difference in the sample means, the authors differ between two tests. One where equal variance is assumed in the two samples and one unequal variance is assumed. Determining whether to assume equal or unequal variance can usually be done by doing an F-test (Bendsen, 2018), which was tested for by the researchers, however, the researchers did conduct both, as the underlying assumptions of the F-test could be wrong. This was not the case, as the same results occurred. In order to carry out the t-test, the authors have used Excel as the primary software. Here, the data-analysis toolpack was used. As illustrated below, the t-test is done by choosing the two sample means as *Variable 1* and *Variable 2* and a significance level of 95%, indicated by the Alpha = 0,05.



When the t-test is done, the authors will conclude of the one-sided P-value as the assumptions of the hypothesis can be stated as follows:

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

Notice that μ refers to the population means and, in this case, the test is on whether one mean is greater than the other. Hence, a one-sided test in conducted. In the case of testing whether the two means is different, a two-sided tale could be developed instead.

Both the proportion test and the t-test assume normality and independence. The normality can be justified if the samples are reasonable large. The independence can be difficult to check and one must have information about the sampling procedure. They also assume representativeness, as random selection method has been used to extract the samples.

The test statistic for proportions is calculated as follows:

$$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Under the null hypothesis $H_0: p_1 = p_2$ this follows a standard normal distribution. The sample proportion follows a normal distribution according to the theory of the central limit theorem given the sample sizes are large enough (Routledge, 2018). When conducting this test, we also assume that the samples are representative and randomly selected.

To carry out the proportion test, the authors have created a binary outcome. Here, the value is 1 if the participants invest the most amount of points in the investigated Company. Otherwise, 0. By doing so the probability that participants picked the best company can be measured. To test whether this is significant, the P-value is calculated from the test statistic.

Calibration

To assess whether experts and non-experts are either well calibrated, the authors will develop a calibration curve. A calibration curve expresses the relationship of accuracy between the participants objective results of investments and their subjective confidence level believed to reflect this investment (Keren, 1987). To develop this calibration curve, the authors have converted confidence levels into intervals. The intervals are 1-3 (weak confidence), 4-5 (medium-low), 6-7 (medium-high) and 8-10 (High confidence). The researchers have chosen to develop intervals to make sure that data points are present within each interval. In addition, a histogram will be presented to visualize the distribution of predictions within experts and non-experts. To test for significance for overconfidence, the authors will reject or confirm the hypothesis through the probability test.

5.9 Case Presentation

The following section introduces the four startups that are presented in the questionnaire, all of which can be found in Appendix 6. The authors have received the cases from Almi Invest – the largest Swedish venture capital and among the best performing venture capitals in Europe (Pitchbook, 2018). The primary reason to present Swedish companies from a Swedish venture capital is to eliminate risk of participants having prior knowledge of the start-ups and hereby decrease bias. The researchers chose historical cases, where the return of investment was known, in order to measure the performance of the crowds. Moreover, the four startups are B2B, meaning their products are sold from business to business. In addition, they are all considered to operate in the Tech industry. As the startups have all gotten early-stage investment, the underlying assumption to the experiment is that the market reach for each startup has a billion-dollar potential. The market potential is an important factor for judging early stage companies (Rea, 1989) – however, this element is removed from the case presentations. The crowd participants are therefore asked to rate the startups in terms of their team and product, which are among the most important factors that VC's evaluate investments on (Franke, Gruber, Harhoff, & Henkel, 2008). Other key features such as the number of employees and the identify of prior investors are also included in the case presentations. Obviously, information about the name of the company and founders are hidden from the cases to eliminate the possibility of participants searching for the case companies on the internet.

All information on the four companies is obtained directly and gathered from the homepage of the companies, articles, databases; Crunchbase, Pitchbook, Dealroom, Linkedin and information directly from Almi Invest.

Neo4J – Company1:

The first case company is Neo4J, which is the startup, that gave a 10x return of investment and therefore is the best investment. Neo4J is a developer of an open-source graph database. It is designed to help all types of companies to understand connections, influences and relationships in data through application that can adapt to changing business environment. The product is focused on visualizing data and relations though graphs and not rows and columns, making it very user friendly to the users (Neo4j, 2018).

The vision of Neo4J is: "*Whether we want to understand relationships between customers, elements in a telephone or data center network, entertainment producers of highly connected data will be key in determining which companies outperform their competitors over the coming decade"* (Neo4j, 2018).

The two founders went to the same university and afterwards joined the same company. They came up with the initial idea in 2001, as they saw ways of improving relational databases. The first prototype was built the year after, in 2002 which gave them the possibility to found the company in 2007.

The company received its first SEED funding of $2.5 million in 2009 from Sunstone Capital and Conor Venture Partner. In 2011, they received their investment from Almi Invest which made it possible for them to move its headquarters to Silicon Valley. The team behind consisted of 10-15 employees at the time of investment.

Dapresy – Company2

Dapresy is the second case company and, provided Almi Invest with a 3x return of investment. Dapresy is a niche company with deep expertise in the market research sector. Their mission is to be a global leader in market research reporting. Dapresy offers its users a product of market intelligence and an insight software. It allows its users to create and deploy visually compelling KPI dashboards for visual business intelligence, marketing research, and customer experience management (Dapresy, 2018). The company's software visualizes and distributes data into existing company processes; enables clients to deploy KPI-driven (key performance indicators) marketing dashboards; and communicates complex data from markets, users, and customers (Dapresy, 2018). Dapresy makes it easy to import raw survey data from leading survey platforms and works with all major formats; SPSS, Triple-S, Excel and non-survey marketing data.

The CEO and co-founder Torbjorn Andersson founded the company in 2003 and his knowledge and passion for technology and marketing research resulted in the birth of the Dapresy. His entire professional career has been dedicated to data visualization. He is a recognized expert in the field and has helped hundreds of companies with innovative visualization processes (Dapresy, 2018).

Telcred – Company3

Telcred, the third case company, is considered the worst investment as it is currently permanently closed and has not provided any return of investment. The company was a spin-off from Swedish Institute of computer science and was founded in 2009. Telcred was a developer of physical access control system whose feature was that its lock controllers did not need to be updated when users' or users' access rights change. Time-limited "digital keys" were sent to an NFC (near field communication) phone or a contactless smart card, which could then be used to gain physical access (Telcred, 2018). Hence, people are able to open doors through the use of their mobile devices or smart cards. Telcred primarily targeted infrastructure companies such as telecom operators, power companies and operators of public transportation (Telcred, 2018).

Before the investment from Almi Invest, Telcred had four employees and had prior investments from Stockholm Innovation & growth AB and STING, which is a Swedish accelerator program.

The CEO and co-founder Carlo Pompili has a history within startups, which is usually a good indicator of success. He has worked with new business creation within the industry of IT and telecom for more than 15 years.

Limes Audio – Company4

The fourth and final startup – Limes Audio, was a great investment for Almi Invest as it provided an 8x return of investment. The company was acquired by Google in 2017 after a lifetime of 10 years. Limes Audio grew out of an academic research at Sweden's Blekinge Institute of Technology. The main product of Limes Audio is an audio software known as TrueVoice (Pitchbook, 2018). It promises to enhance quality voice communication through a combination of acoustic echo cancellation, noise reduction and automated mixing, among other tweaks (Pitchbook, 2018). The technology and service of Limes Audio gained very early traction world-wide from some of the most renowned companies in the audio business (Pitchbook, 2018). The reason is the technology benchmark results, a proven track record, a flexible solution and a dedicated team (Pitchbook, 2018). The team consisted of 6 employees in total, where the two founders had prior experience within the industry and both pursued a Ph.D. in Applied Signal Processing.

Chapter 6 - Findings

To answer the research question*; How can crowds' predictions be used to optimize the low success rate of Venture Capital and are predictions victim of overconfidence?* – and furthermore, to test the 6 hypotheses, the data will be presented, interpreted and discussed.
Following section will first of all only present the findings. The following chapters will include an interpretation and discussion. The findings will provide an executive summary of the main findings. Secondly, the authors will present information of the overall findings in terms of descriptive statistics and relevant correlations. Lastly, the findings of each hypothesis will be treated independently followed by a brief conclusion. The full Excel file regarding calculations and tables can be found in Appendix 8.

6.1 Executive Summary of Findings

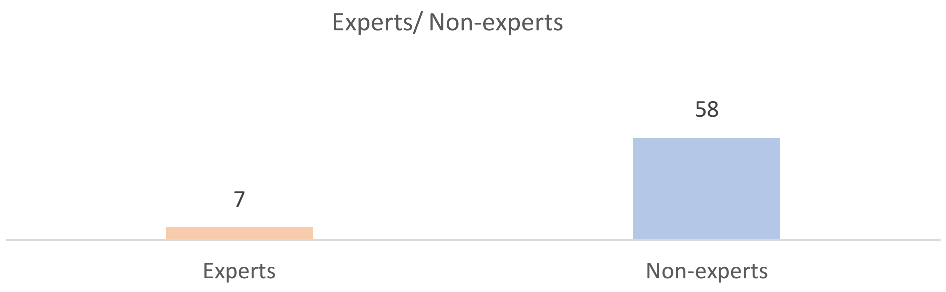
The main findings to the study is that both the total crowd and the non-expert crowd managed to outperform the experts as they generated the most points. Though, the difference in sample means were insignificant. Moreover, the foreign experts and experienced crowd outperformed Danish experts and less experienced experts. However, due to the very small sample size of experts, which also showed insignificance, it is difficult to derive conclusions on which expert crowd were the best forecasters. The last two hypotheses also showed insignificance. Though, findings suggest that experts are generally better calibrated than non-experts, however, the experts - on average showed overconfidence compared to non-experts.
A significant finding was the correlation between Likert rating and the amount invested which also ensures internal reliability. In addition, the non-expert crowd showed that the Product were the primary driver behind their investment, whereas experts, on the contrary, expressed a significant correlation between team and investment.

6.2 Overall findings

Sample demographics

As presented in the literature review, a crowd is able to make wise decisions in the existence of diversity, independence and decentralization (Surowiecki, 2004; Hong & Page, 2001;2004). The total crowd in this experiment consisted of 65 participants which was further divided into two major crowds; *Experts and Non-experts.* As Table 1 shows, the expert group consisted of 7 participants experts who are working in a Venture Capital, which of whom were from Denmark and London.

Table 1



Despite the limitations of the research, which will be further elaborated on later in the report, the research strived to impose diversity in terms of gender, age, nationality, field of expertise, educational background and employment status. The variables containing the most diversity will therefore be presented below.

Table 2, 3 and 4 presents the distribution of these variables.

Table 2

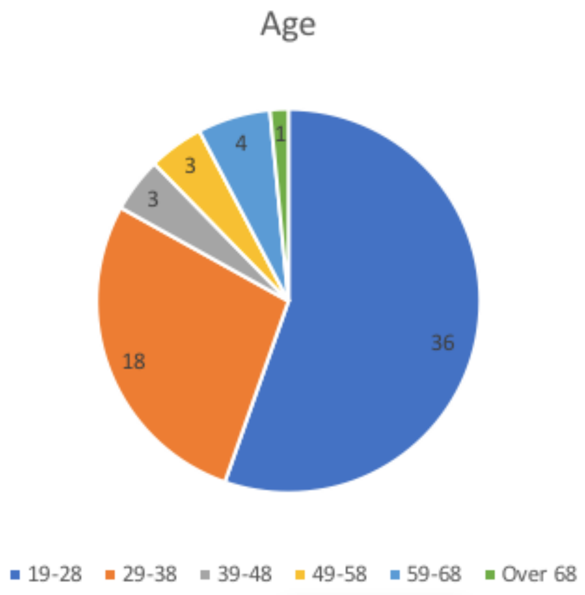


Table 3

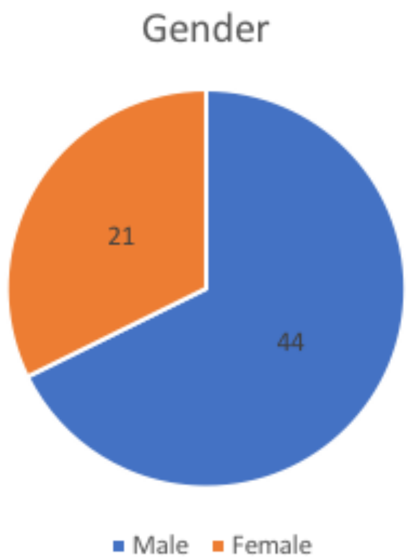
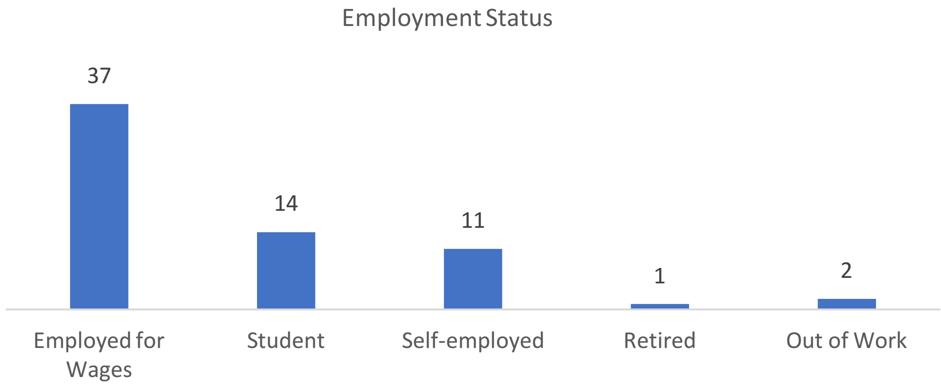


Table 4



The Educational Background of the crowd participants consisted of 58,2% having a Master Degree, 26,9% Bachelor Degree, 9% High School Graduates, 4,5% Doctor Degree and 1,5% Trainee. Moreover, 55 of the participants in the sample was from Denmark, whereas as the remaining was spread into the nationality of France, England, Iceland, Slovakia, Sweden, Finland and Ukrainian. Additionally, the crowd participants had a diverse background in terms of field of expertise. 18 of the total participants had experience as an entrepreneur. Other fields of expertise ranged from law, leadership, strategy, marketing, management, public relations, sales, logistics, advertising, politics, medicine, phycology, IT, journalism and sports.. The full review on the sample demographics can be found in Appendix 1.

Company Evaluation

Table 5 shows the average rating each company has received on the 5-point Likert scale. Important to note, is that the average is across the total crowd participants. Furthermore, it shows that Company 4 received the highest rating, Company 1 the second highest, Company 3 the third highest and lastly, Company 2 with the lowest rating. In other words, the crowd perceived Company 4 and 1 as having the highest quality in terms of product and team.

Table 5

Likert Score - Total Crowd	Company 1	Company 2	Company 3	Company 4
Average	3,52	3,32	3,46	3,72

Table 6 below, shown the average investment that each company has received. Once again, Company 4 got the highest amount of points on average, Company 1 second highest, Company 3 third, and Company 2 the lowest amount of points. This again, indicates that by simply aggregating the judgements across the crowd participants, the crowd invested the most points on average in Company 4 and 1, which are the two start-ups with the highest return of investment.

Table 6

Invested points - Total Crowd	**Company 1**	**Company 2**	**Company 3**	**Company 4**	**Total Return**
Average	241,69	216,92	234,38	307,00	5.523,69

Demographic Correlations

In order to investigate which individuals or groups that were the best performing, in-depth tests for correlations between demographic variables of the sample was conducted. Here the researchers found little and no evidence of correlations between performance, in terms of total return of investment and demographic variables such as gender, age, educational background, employment status or whether you were an expert within VC's or not. These correlation matrixes can be found in Appendix 3. However, females did, on average, generate a marginal higher return of investment than males. Females is the sample generated 5772 points and males, 5521 points.

Rating vs Investment

Another interesting finding is the correlations between how the start-up was rated in terms of quality and the number of invested points each company received. As shown in Table 7 below, the numbers of the Likert score that each company has received, in terms of team, product and in total, correlates with the amount of invested points. It gives the authors a clear indicator that the Product of the startups had the highest impact for how much points participants choose to invest. In addition, this finding also increases internal reliability of the experiment.

Table 7

TOTAL CROWD

Correlations		**Company 1**	
Score vs Investmen	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,62	0,50	0,68
R-Square	**0,38**	**0,25**	0,49
Adj. R-Square	0,37	0,23	**0,47**
Significance	0,00	0,00	0,00

Correlations		**Company 2**	
Score vs Investmen	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,60	0,44	0,62
R-Square	**0,37**	**0,20**	0,41
Adj. R-Square	0,36	0,18	**0,39**
Significance	0,00	0,00	0,00

Correlations		**Company 3**	
Score vs Investmen	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,71	0,50	0,72
R-Square	**0,52**	**0,25**	0,53
Adj. R-Square	0,51	0,24	**0,52**
Significance	0,00	0,00	0,00

Correlations		**Company 4**	
Score vs Investmen	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,62	0,11	0,62
R-Square	**0,39**	**0,01**	0,39
Adj. R-Square	0,38	0,00	**0,37**
Significance	0,00	0,36	0,00

Chin (1998) identified how to analyse certain levels of adjusted R-square. He considered 0.67, 0.33 and 0.19 to be substantial, moderate and weak, respectively. Meaning that all findings on the adjusted R-square values, in this experiment, provides moderate proves of the relationship between the model and the response variables.

To determine whether the correlations between the Likert rating and the invested points is generated by chance or if the relationship is considered significant, the authors conducted both a multiple and simple regression analysis (Appendix 8). The team of Company 4 is considered to have no correlation to invested point, which is also the only correlation that showed insignificance. All other correlations between, product, team, total and invested points is significant.

Law of Large Numbers

The last and most interesting finding is that by increasing the size of the total crowd and thereby also diversity, the total return of investment across all four startups also increased. As the theory of states (Surowiecki, 2004; Keuschnigg & Ganser, 2017; Hong & Page, 2001;2004) as well as the general Law of numbers (Routledge, 2018), increasing size and diversity of the crowd, will eventually result in a more accurate answer. In this case, the crowd were able to generate more points, by investing in the best startups, as the researchers increased diversity and sample size. This was tested by dividing the total crowd into three groups by the timestamp of participation. This can also be explained as N=21, N=42 and N=65. As shown below in Table 8, the total average return of investment increased. Even though, none of the results showed significance at a 95% confidence interval, the trend of the P-value was positive by an increase in N. At N=21 the P-value was 0.43 and at N=65 the P-value was 0.24.

Table 8

N	Avg. Return of Investment	P-Value
21	4932,14	0,43
42	5163,72	0,43
65	5523,69	0,24

6.3 Hypothesis testing

Following seeks to introduce main findings in regards to the hypotheses. First, the most important descriptive statistics will be presented, relevant correlations between variables and lastly a student's t-test and a probability test. The descriptive statistics of how each company was rated and invest in, is presented in Appendix 2. The most important findings have been drawn from these to be presented. Furthermore, the investigated hypothesis will be mentioned as **H1, H2, H3** etc. during the findings and discussion.

An important note is that "Accuracy", in terms of the crowd performance, is measured through their total return of investment in the task of investing 1.000 points.

H1: The total crowds' prediction is more accurate than the expert crowds' prediction

H2: The crowd of non-experts are more accurate than the expert crowds in their prediction

H3: Foreign experts are better predictors than Danish experts

H4: Experienced venture capitalists are more accurate in their prediction than their less experienced venture capitalists.
H5: Experts are better calibrated than non-experts
H6: Both experts and non-experts are overconfident in their predictions, however experts are relatively more overconfident than novices.

Hypothesis 1

In terms of **H1**, the total crowd consist of both experts and non-experts. The purpose is to test whether the total crowd can outperform the experts (Venture Capitalists).

As table 9 shows below, the Total Crowd does in fact outperform the experts with total return values of 5523.69 against 5035.71. In terms of the order of which companies that has received the most points on average, the total crowd managed rank Company 4 and Company 1, as the best and second best, respectively. On the contrary, the Experts invest the most in Company 4 and Company 3. As Company 3 is the least profitable startup, this is arguably the reson to the lower value in Total Return. Company 1, the best startup of the study, receives the lowest points on average by the Experts whereas the Total Crowd invest second most in that company.

Table 9

Invested points - Total Crowd	**Company 1**	**Company 2**	**Company 3**	**Company 4**	**Total Return**
Average	241,69	216,92	234,38	307,00	5.523,69
Variance	26.271,31	21.056,79	34.013,68	42.450,63	3.072.583,81
Std. Dev.	162,08	145,11	184,43	206,04	1.752,88

Invested points - Experts	**Company 1**	**Company 2**	**Company 3**	**Company 4**	**Total Return**
Average	164,29	178,57	300,00	357,14	5.035,71
Variance	20.595,24	11.547,62	25.833,33	30.357,14	3.030.595,24
Std. Dev.	143,51	107,46	160,73	174,23	1.740,86

** Each value in the table is that average points that each company has received out of the total of 1.000. The Total Return value is the number of generated points. This is calculated by multiplying the points with the return of investment of the specific startup. Company 1: 10x – Company 2: 3x – Company 3: 0x – Company 8x.*

Despite the results shown above, the authors conducted a t-test to test whether the sample means are significantly different from each other. Here, as visualized in Table 10, the results showed insignificance with a P-value >0.05. Therefore, the researchers can reject H1, as there is evidence that the total crowd is more accurate than the expert crowd.

Table 10

T-test - Equal Variance	Total Crowd	Experts
Mean	**5523,69**	**5035,71**
Variance	3072583,81	3030595,24
Observations	65	7
Pooled Variance	3068984,79	
Hypothesized Mean Difference	0	
df	70	
t Stat	0,70	
P(T<=t) one-tail	**0,24**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,49	
t Critical two-tail	1,99	

As Company 1 is the best investment, it is interesting to investigate which whether the difference in the mean, in terms of invested points in Company 1, is significantly different. The mean showed that the total crowd generated 2416,92 points on average whereas the Experts managed only 1642,86 on average (Table 11). To test for significance, the same approach as before was done and the P-value showed 0.11. As this value is >0.05, it is considered insignificant. Moreover, the probability test, which shows how many the percentage of the crowd were able to pick the best investments, 25% of the total crowd participants picked Company 1 – meaning that they invested the most point in that startup (Appendix 8).

Table 11

Company 1		
T-test - Equal Variance	**Total Crowd**	**Experts**
Mean	**2416,92**	**1642,86**
Variance	2627131,01	2059523,81
Observations	65	7
Pooled Variance	2578478,964	
Hypothesized Mean Difference	0	
df	70	
t Stat	1,21	
P(T<=t) one-tail	**0,11**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,23	
t Critical two-tail	1,99	

The results on how the crowd performed in terms of investing in Company 4, showed that the Total crowd generated 2456 points on average from their investments, whereas the Expert crowd generated 2857,14. However, the P-value, as shown in Table 12, of these two means is 0,27 – meaning the difference is insignificant. The categorical test showed that 32% of the total crowd invested most points in Company 4 (Appendix 8).

Table 12

Company 4		
T-test - Equal Variance	**Total Crowd**	**Experts**
Mean	**2456,00**	**2857,14**
Variance	2716840,00	1942857,14
Observations	65	7
Pooled Variance	2650498,612	
Hypothesized Mean Difference	0	
df	70	
t Stat	-0,62	
P(T<=t) one-tail	**0,27**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,54	
t Critical two-tail	1,99	

Hypothesis 2

This following section will present the results of **H2**. This hypothesis focusses solely on the relationship between experts and non-experts. Experts consisting of a sample of N=7 and non-experts of a sample of N=58.

Table 13

Invested points - Experts	**Company 1**	**Company 2**	**Company 3**	**Company 4**	**Total Return**
Average	164,29	178,57	300,00	357,14	5.035,71
Variance	20.595,24	11.547,62	25.833,33	30.357,14	3.030.595,24
Std. Dev.	143,51	107,46	160,73	174,23	1.740,86

Invested points - Non-experts	**Company 1**	**Company 2**	**Company 3**	**Company 4**	**Total Return**
Average	251,03	221,55	226,47	300,95	5.582,59
Variance	26.505,05	22.224,74	34.878,95	44.122,33	3.098.136,18
Std. Dev.	162,80	149,08	186,76	210,05	1.760,15

** Each value in the table is that average points that each company has received out of the total of 1.000. The Total Return value is the number of generated points. This is calculated by multiplying the points with the return of investment of the specific startup. Company 1: 10x – Company 2: 3x – Company 3: 0x – Company 8x.*

The average is the main indicator for testing **H2**. The results show, in Table 13, that experts on average gain a total return of 5035,71 points. Whereas, non-experts gain an average of 5582,59 points in total return. From the findings, it is seen that experts on average invest 164,29 points in Company 1 and 357,14 points in Company 4. Conversely, non-experts invested an average of 251,03 points in Company 1 and 300,95 points in Company 4. Hence, experts invest generally on average less points in the companies considered "good investments". In terms of ranking the non-experts manages to rank Company 1 and Company 4 as the investments they respectively invest most in on average. While experts on average rank Company 3 and Company 4 as the two "best" investments. Both experts and non-experts invest on average most points in Company 4. Non-experts invest on average the least amount of points in Company 2, whereas experts invest the least point in Company 1. Experts invest on average second most in Company 3, which is the least profitable start-up, with a value of 300. Non-experts invested on average the second least points in Company 3 with a value 226,47 points on average.

In order to test for significance of the findings of average total return between experts and non-experts a t-test assuming equal variance was conducted. Equal variances are assumed as a f-test prior to the t-test was conducted, that confirmed that this assumption. The findings were insignificant as the p-value were > than 0,05. Hence, the **H2** is rejected, as there is no evidence that the crowd of non-experts are better than the expert crowd.

Table 14

T-test - Equal Variance	**Non-Experts**	**Experts**
Mean	**5582,59**	**5035,71**
Variance	3098136,18	3030595,24
Observations	58	7
Pooled Variance	3091703,706	
Hypothesized Mean Difference	0	
df	63	
t Stat	0,78	
P(T<=t) one-tail	**0,22**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,44	
t Critical two-tail	2,00	

In order to test for whether the non-experts significantly invested more points in Company 1, that gave the highest return of investment, a t-test was conducted on the two means. The same procedure as mentioned above for testing for equal variances was conducted. However, the results were not significant as shown in Table 15.

Table 15

Company 1		
T-test - Equal Variance	**Non-experts**	**Experts**
Mean	**2510,34**	**1642,86**
Variance	2650505,14	2059523,81
Observations	58	7
Pooled Variance	2594221,206	
Hypothesized Mean Difference	0	
df	63	
t Stat	1,35	
P(T<=t) one-tail	**0,09**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,18	
t Critical two-tail	2,00	

The above t-test was conducted to test whether non-experts significantly rated Company 1 a higher total score on the Likert-scale. The results were significant as shown in Table 16, however, the **H2** can is still rejected. The total Likert-score is generated from the participants ratings of "product" and "team" on the Likert-score. Additionally, as non-experts were shown to invest more points in Company 1 on average, the researchers found it interesting to see the relation between this and the Likert-scale.

Table 16

Company 1		
T-test - Equal Variance	**'otal Crowd**	**Experts**
Mean	**3,57**	**3,14**
Variance	0,42	0,31
Observations	58	7
Pooled Variance	0,4060521	
Hypothesized Mean Difference	0	
df	63	
t Stat	1,67	
P(T<=t) one-tail	**0,05**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,10	
t Critical two-tail	2,00	

In alignment with the abovementioned test for investments in Company 1, the same test was conducted for Company 4, as these two are considered as the “good investments”. The findings demonstrate that experts on average invest more in Company 4 than non-experts. Nonetheless, the results were found not to be insignificant in both the investment and Likert rating task.

Table 17

Company 4

T-test - Equal Variance	**Non-experts**	**Experts**
Mean	**2407,59**	**2857,14**
Variance	2823829,16	1942857,14
Observations	58	7
Pooled Variance	2739927,062	
Hypothesized Mean Difference	0	
df	63	
t Stat	-0,68	
P(T<=t) one-tail	**0,25**	
t Critical one-tail	1,67	
P(T<=t) two-tail	0,50	
t Critical two-tail	2,00	

In order to give the findings a higher degree of robustness, the probability of participants who invested in the two “good” investments were tested for. This is an additional finding, that can support the findings of the mean already accounted for above. The results show that both in Company 1 and Company 4 the non-experts invested in these with a higher probability. The highest variance is seen in invested points in Company 1, where 26% of the non-experts invests the most and only 14% of the experts. This corresponds well with the findings, that non-experts generally invest more in Company 1 on average. Additionally, the results show that more participants in general have invested Company 4 than Company 1. The results demonstrate, that 29% of experts invest the most amount of points in Company 4 compared to 33% non-experts. The finding shows that less experts invest in Company 4 in terms of percentage compared to non-experts, which could seem conflicting to the results that experts on average invest more in Company 4. However, this show that the probability of experts that actually invest the most in Company 4, invest a relative higher amount in Company 4 than non-experts. The two probability tests were however insignificant as shown in Table 18.

Table 18

	Invested most in Company 1	
	Experts	Non-experts
Yes	1	15
No	6	43
%	14	26
Number	7	58
P-value	0,749	

	Invested most in Company 4	
	Experts	Non-experts
Yes	2	19
No	5	39
%	29	33
Number	7	58
P-value	0,589	

* *To determine whether experts invested the most in either Company 1 or 4, the authors created a binary outcome, which is set to be true if the invested points in the specific company were of the highest value compared to the others.*

The researchers investigated whether correlations were found between the Likert-scale and the investments respectively made by non-experts and experts. To demonstrate if the underlying factors in Likert-scale of "Product" and "Team" had any specific correlations to the investments a multiple regression-analysis was conducted. The results, as presented in Table 19 and 20, showed that for non-experts the "product" were the main indicator for their investment, as the highest correlations were found here. For the non-experts, it was found to be significant in case of all the start-ups. Looking at the expert's rating of the "Product", this showed a higher correlation compared to the "Team" in Company 1, Company 3 and Company 4. Thus, the correlation shows that "Product" accounted the most for the expert's investments in these. However, correlations showed that expert's investments had a higher correlation to the "Team" of Company 2 than the "Product". This finding was significant. Hence, expert's investments are to a small degree based more on the "team" than the non-experts investments. Significance levels of the variable accountable for non-experts and expert's investments are shown in the Table 19 and 20 below.

Table 19

EXPERTS

Correlations		**Company 1**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,72	0,07	0,73
R-Square	**0,52**	**0,01**	0,53
Adj. R-Square	0,43	-0,19	**0,29**
Significance	0,07	0,88	0,22

Correlations		**Company 2**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,70	0,85	0,91
R-Square	**0,49**	**0,72**	0,83
Adj. R-Square	0,38	0,67	**0,75**
Significance	0,08	0,02	0,03

Correlations		**Company 3**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,53	0,48	0,61
R-Square	**0,28**	**0,23**	0,37
Adj. R-Square	0,14	0,07	**0,05**
Significance	0,22	0,28	0,40

Correlations		**Company 4**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,77	0,20	0,84
R-Square	**0,59**	**0,04**	0,71
Adj. R-Square	0,51	-0,15	**0,57**
Significance	0,04	0,67	0,08

Table 20

NON-EXPERTS			
Correlations		**Company 1**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,60	0,51	0,68
R-Square	**0,36**	**0,26**	0,48
Adj. R-Square	0,35	0,25	**0,46**
Significance	0,00	0,00	0,00
Correlations		**Company 2**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,60	0,41	0,62
R-Square	**0,36**	**0,17**	0,39
Adj. R-Square	0,34	0,15	**0,37**
Significance	0,00	0,00	0,00
Correlations		**Company 3**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,71	0,50	0,72
R-Square	**0,55**	**0,26**	0,56
Adj. R-Square	0,51	0,24	**0,52**
Significance	0,00	0,00	0,00
Correlations		**Company 4**	
Score vs Investment	**PRODUCT**	**TEAM**	**TOTAL**
Multiple R	0,62	0,11	0,62
R-Square	**0,39**	**0,02**	0,39
Adj. R-Square	0,38	0,00	**0,37**
Significance	0,00	0,36	0,00

Hypothesis 3

H3 states that foreign experts are better at predicting future startups than Danish experts. Participants are considered experts by having experience within Venture Capital. The foreign crowd consist of three experts working in a London Venture Capital. All of which are from different levels of management and experience. The Danish experts are also a mix between experience and management levels in different VC's.

The results regarding **H3**, is presented in Table 21. The most obvious finding is that Foreign Experts outperforms Danish Experts in terms of the sample means. In other words, the foreign Experts managed to generated higher total average return, which might be a reason that the foreign crowd invested more, on average, in Company 1. However, the P value of 0.38 showed that the difference in means is insignificant and thereby the **H3** is rejected.

Table 21

T-test - Equal Variance	Danish Experts	Foreign Experts
Mean	**4837,50**	**5300,00**
Variance	3732291,67	3310000,00
Observations	4	3
Pooled Variance	3563375	
Hypothesized Mean Difference	0	
df	5	
t Stat	-0,32	
P(T<=t) one-tail	**0,38**	
t Critical one-tail	2,02	
P(T<=t) two-tail	0,76	
t Critical two-tail	2,57	

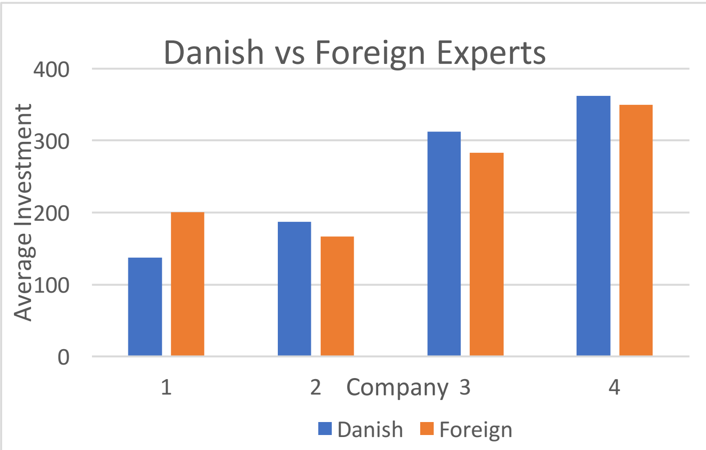


Figure 6

Hypothesis 4

The purpose of **H4** is to investigate whether experienced experts are more accurate in their prediction than less experienced experts. To distinguish between experienced and less experienced experts, the researchers grouped experienced experts at a level of >10 years of experience. Here the two crowds consisted of 4 less experienced experts and 3 experienced.

Table 22

T-test - Equal Variance	Experienced Experts	Less Experienced Experts
Middelværdi	**5166,67**	**4937,50**
Varians	4203333,33	3228958,33
Observationer	3	4
Puljevarians	3618708,333	
Hypotese for forskel i middelværdi	0	
fg	5	
t-stat	0,16	
P(T<=t) en-halet	**0,44**	
t-kritisk en-halet	2,02	
P(T<=t) to-halet	0,88	
t-kritisk to-halet	2,57	

As shown in Table 22, the mean of the two samples indicates that experienced experts managed to generate 5166,67 points on average. The less experienced produced 4937,5 points on average by investing in the four startups. This might be caused by the less experienced experts invested more in Company 3 – as shown in Figure 7, which is the least profitable. To test whether the difference in means were significant, the t-test showed a P-value of 0,44. The difference is therefore insignificant and the authors can reject **H4**, meaning there is no evidence that experienced experts are more accurate in predicting start-ups.

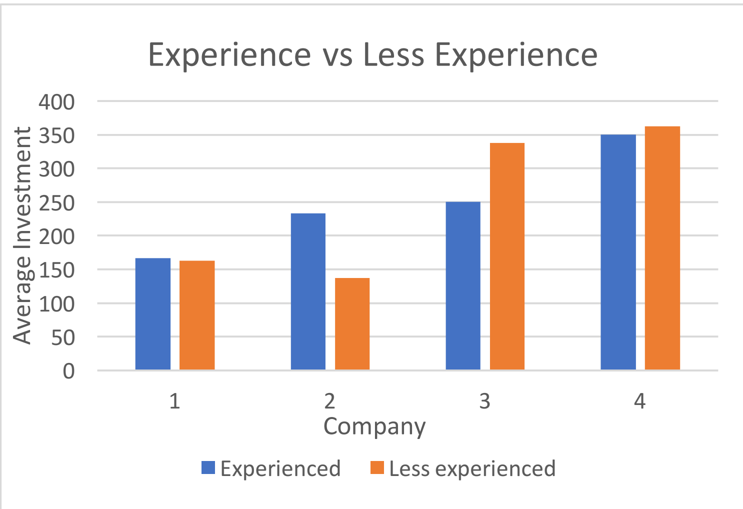


Figure 7

Hypothesis 5

H5: The following section is going to present the results for H5. It will solely focus on whether experts are better calibrated than non-experts and places no emphasize on how accurately the respective groups are calibrated. This will be accounted for later on in H6.

Table 23

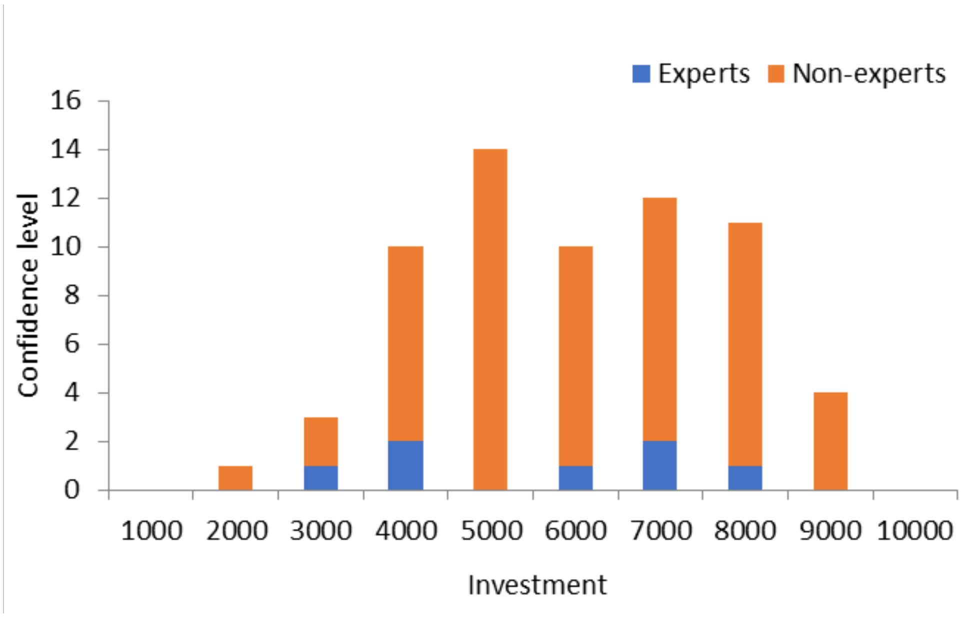
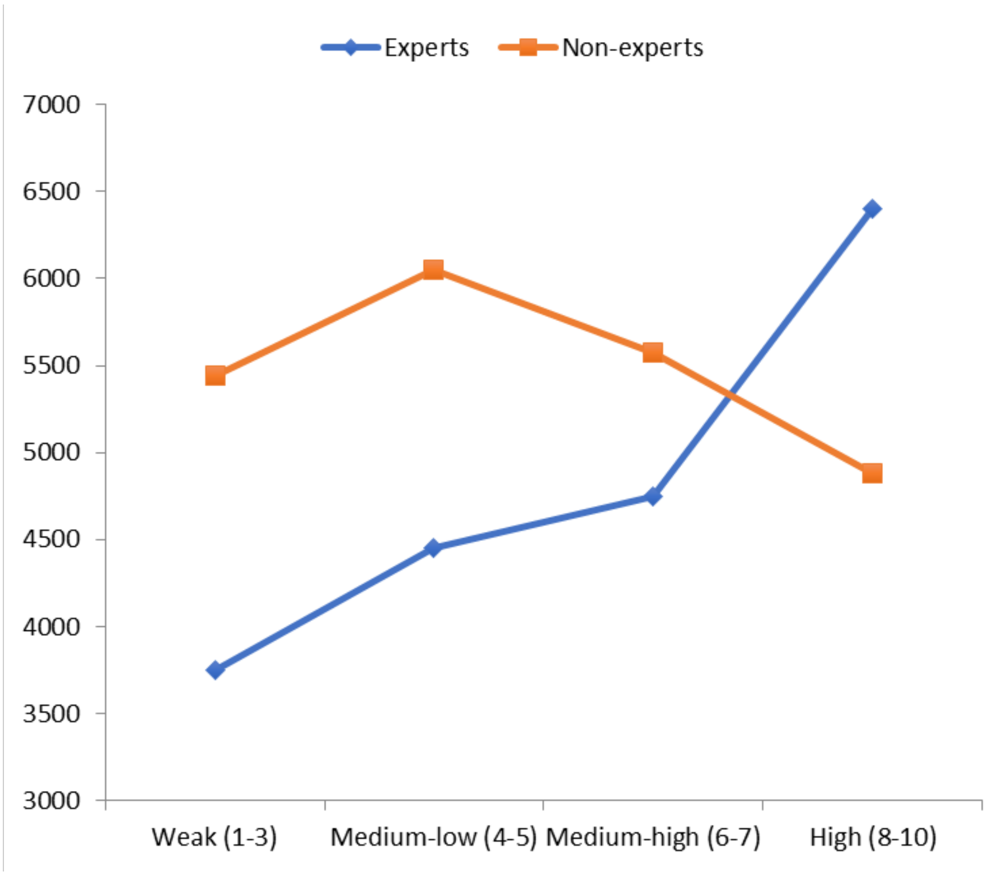


Table 23 shows a roughly normal distribution of the average return of experts and non-experts investments in points. The minimum return of investment generated 1950 points and the maximum produced 9000 points (Appendix 2). The average of the investments were 5524 points. In order to test whether experts were better calibrated than non-experts the findings from table 23 were cross matched with the individual's respective confidence level, that they had indicated through the survey. As seen in table 24, the curve of experts shows a better fit of expert's confidence level and their investments. Meaning that experts show more consistency between their investments and the confidence they rate this investment by. Additionally, to table 24 a regression analysis was conducted to test the significance of the correlations. The analysis showed a correlation between expert's investments and their confidence level of 0,61. Non-experts had a correlation of -0,04 between their investments and their confidence level. Thus, experts fit between investments and confidence level are more accurate than that of non-experts. However, the results are not significant and the **H5**, stating that experts are better calibrated than non-experts, is therefore rejected.

Table 24



Hypothesis 6

H6 states that both experts and non-experts are overconfident in their predictions. Moreover, experts are relatively more overconfident than non-experts. Table 25 is a descriptive table across the total crowd in terms of their total return of investment and confidence level.

Table 25

TOTAL RETURN		CONFIDENCE LEVEL	
Middelværdi	5523,69	Middelværdi	5,31
Standardfejl	217,4178469	Standardfejl	0,275544
Skævhed	0,127246415	Skævhed	-0,11094
Minimum	1950	Minimum	1
Maksimum	9000	Maksimum	10
Antal	65,00	Antal	65,00

Table 26 shows the average between Experts and non-experts, in terms of total return across the crowds and the average confidence level. First and most noticeable is, by looking at the relationships between the average Total return and confidence level, the findings impose that Experts are relatively more overconfident than Non-experts. Non-experts, on average, showed a slight under confidence in their predictions. However, none of these findings was significant with P-values >0,05.

Table 26

Average Confidence vs Total Return	Total Return	Confidence
Experts	5035,71	6,00
Non Experts	5582,59	5,22
Total Crowd	5523,69	5,31

Additionally, a test for proportions was conducted. Overconfidence was in this test measured as if the participants generated a higher return than their level of confidence. As visualized below, 43% of the experts showed overconfidence in their predictions. On the contrary, only 29% of Non-experts revealed overconfidence. Despite the P-value of 0,232 and therefore insignificant, the results were aligned with the investigated hypothesis. As the P-value, shown in Table 27, conducted through the probability test were insignificant (Appendix 8) the **H6** is rejected by the authors and therefore it is not possible to state that both experts are relatively more overconfident in their predictions than non-experts.

Table 27

	Overconfidence	
	Experts	**Non-experts**
Yes	3	17
No	4	41
%	0,43	0,29
Number	7	58
P-value	0,232	

Chapter 7 – Discussion

The objective of the thesis is to investigate whether *The Wisdom of Crowds* can be deployed within venture capital to create more accurate predictions of start-ups' success. Moreover, the thesis investigates if certain characteristics of venture capitalists affect the accuracy of predictions. Finally, as overconfidence is "*regarded as one of the most prevalent judgment biases"* (Glaser, Langer & Weber, 2012, p.1) and is shown in many studies to lead poorly decisions (Andersson, Edman, & Ekman, 2005; Kahneman & Tversky, 1973) it is investigated if overconfidence is present within venture capital decisions making, which could be expected due to the high frequency of investments that end up with not generating the expected return. To the best of the researcher's knowledge, no others have tested this within the industry of Venture Capital.

The following section will discuss the findings of the study and relate them to previous research. First, the investigated hypotheses will be elaborated on – secondly, a general discussion of the findings will be presented.

7.1 Hypotheses discussion

Is Combining knowledge the best?

H1 was tested in order to see whether the total crowd was significantly better than the expert crowd. Crowds that are diverse and independent have proved to be able to produce the very accurate predictions (Surowiecki, 2004; Keuschnigg & Ganser, 2017). The primary reason for this hypothesis is to see whether a crowd consisting of knowledge from experts and non-experts will beat a group of experts. Theory states that the larger the size of the crowd, the more accurate the prediction (Keuschnigg and Ganser, 2017). Studies on the phenomenon of The Wisdom of Crowds have shown that it is capable of ranking certain items (Lee et al., 2012; Grazing, 2016).

- The findings from **H1** showed that the total crowd were able to outperform the experts. It showed that the mean of total crowd was close to the truth (the return of investments). However, the results were insignificant.
- The total crowd invested the most point in Company 4 and Company 1 on average, as the first and second place respectively. The exact order was not consistent with the latent truth - however, the authors separated the four companies into two 'good' investments and two 'bad' investments.

- The categorical calculation was done to reach more robust conclusions. Additionally, it was done to measure the performance of the total crowd, where the results showed that 25% did 'pick' Company 1 and 32% picked Company 4. In other words, over 57% of the crowd invested the most points in the two best companies

In alignment with the theory of *The Wisdom of Crowds*, (Surowiecki, 2004; Hong & Page, 2001; 2004, Armstrong, 2001) showed by adding diversity, independence to the crowd and furthermore, aggregating human judgments, a more accurate answer presented itself. With this assumption in mind, the total crowd managed to identify the two best investments, whereas the expert crowd, on the contrary, did not manage to do so. Hence, it could be argued that, for venture capitals to enhance their performance, they should include non-experts into their investment process of screening start-ups. As the findings get more accurate by increasing the number of participants the law of large numbers is likely to have an effect within venture capital. Therefore, it could be argued that venture capitals should to their best ability increase the number of people that takes part in the process of screening for the best start-ups. It makes sense that the total crowd are able to identify good start-ups, as many crowd participants are likely to act as end customers and can be an indicator of potential demand for a product. Lastly, to generate significant results, the crowd size of the crowds should be increased, as a positive development of the P-value was found when volume was added to the crowd.

Does Diversity beat Ability?

H2 was a test on whether non-experts were more accurate at predicting start-ups than experts. Non-experts are shown to be just as good, or better, at forecasting future events (Tetlock, 2015; Anderson et al., 2005). As the success rate of VC's are low, the authors argued that this is a result of bias and lack of diversity. Venture Capitals, as mentioned earlier, lacking diversity are in terms of gender and other elements which furthermore, highly affect the portfolio of invested start-ups. Additionally, diversity is shown to outperform ability in large crowds (Keuschnigg & Ganser, 2017). Thus, individuals should be cautious of relying on a few individuals' knowledge and thereby 'Chasing the Expert' (Larrick and Soll, 2006).

- The non-experts clearly outperformed the experts when aggregating the investments and by calculating probability for investing in the best companies. As presented in the findings, the **H2** is rejected as there was no evidence that non-experts were better than experts.
- Non-experts invested on average the most points in the two best investments - Company 1 and Company 4. Contradictory, the experts invested the most in the second best and the worst start-up. In fact, the experts invested the least amount of points in Company 1, which was the best investment in terms of return of investment.
- The non-experts rated Company 1 significantly higher than the experts and therefore they are argued to be more accurate in spotting the best investment.
- 26% of the non-experts invested most point in Company 1 compared to only 14% of the experts.
- One significant correlation of the expert's predictions was found between the team and the investment in company 2.
- The findings showed that there was a significant relationship between the Likert-scale rating and the invested points in the specific start-up.
- The correlation between the team and invested points were stronger for the expert crowd compared to the crowd of non-experts.

The findings indicate that non-experts outperform experts in clear correspondence with Tetlock's (2015) findings. This could indicate that when uncertainty and risk is considered high, as it is within Venture Capital, the prediction task is too complex for few individual experts to cope with, which could be a result of e.g. information effects (Andersson, Edman, & Ekman, 2005). This might affect venture capitalists to led their System1 take over even though they were in need of their System2, which can lead to biased decisions (Kahneman, 2003). Hence, it is argued that VC's could benefit from the use of a non-experts crowd. Otherwise, they should at least consider concentrating on debiasing their decisions and additionally, act more in alignment with the law of large numbers and adding diversity into their own crowd of experts. As the non-experts crowd outperforms the expert crowd this corresponds well with the fact that diversity outperforms ability. Therefore, it is important for VC's to focus on creating the most diverse team as possible when dealing with large groups.

Many might be misled by the idea of aggregating human judgements as it signals 'average' performance. Nobody wants to go with the average performance in critical investment decisions. However, this study has shown that by aggregating across the crowd participants this will in fact lead to a more accurate answer. Studies have shown that when groups fail to perceive the value of averaging, this will eventually harm their accuracy (Larrick and Soll, 2006). As the non-experts were able to identify the two best investments. This is in close correlation with the studies of Klein and Garcia (2015), that proved how non-experts were able to filter away "bad" ideas. Yet, they were in this case also capable of identifying the two best. As VC's screen a large volume of start-ups the usage of a non-expert crowd could argued to be a good method to filter in these investments, by either finding the few best or filtering away those they see as bad start-ups.

Correlations between the "team" and the experts' invested points were found to be significant in terms of one start-up. In other words, experts valued the "team" higher when evaluating the start-ups. As the presentation of the start-ups could show a lack of information, especially regarding the team, it is very interesting that non-experts were able to invest the most in the two best start-ups. Hence, this could act as the reason that experts are not as accurate as non-experts, as they placed more emphasis on the team characteristics in some cases. Thus, venture capitalists might benefit from placing more value on the actual idea and product offering.

Summary

In conclusion to the discussion of H1 and H2, the total crowd, consisting of both experts and non-experts, clearly outperformed the experts. Both in terms of the total average investment and the probability test that indicates the percentage that invested the most in the two best start-ups. This is extremely interesting as it corresponds with the fact that "*groups of diverse problem solvers can outperform groups of high-ability problem solvers*" (Hong & Page, 2004, p. 1). As the crowd of non-experts actually performed better than the total crowd, this shows that in this study, the experts have a negative effect on the overall prediction. Hence, this indicate that diversity beats ability.

Moreover, the significant findings of the correlation between Likert-scale rating and invested points is an indicator that experts to some extent puts more emphasis on the team when judging start-ups. The team is also seen in previous literature as a common way for experts to judge a start-up (Franke et al., 2008). Non-experts, on the contrary, was more product orientated, when evaluating the start-up in the two given tasks. As this is the case, it could be interesting to make the non-experts evaluate several parameters within the product offering of the start-up, as it would provide a more detailed view of the perceived values and benefits. Additionally, the same could be done for the team. Overall, it points to a direction that VC's, through the usage of The Wisdom of Crowds, could increase their ability to spot the highest potential start-ups and therefore dramatically increase their success-rate. Previous research has shown that non-experts are capable of sorting away "bad" ideas (Klein and Garcia, 2015). Though, in this study, the total crowd showed that they can also identify the best investments. In fact, including the experts in the crowd did not have an effect on the results of ranking the four companies.

The insignificant results are primarily a consequence of the relative small sample size. Nonetheless, the number of experts in the crowd reflects reality, as decision-making often is done by a few individuals in VC's. Though, it can be argued that if sample size of both experts and non-experts were larger, the results would have shown significance.

Who are the better forecasters? London vs. Copenhagen

H3 was a test to see whether foreign experts were better at predicting than Danish experts. This hypothesis was derived from a global perspective of thought, where a clear difference in performance of Ecosystems is present. London experts are considered better than Danish experts at judging start-ups as the performance of the two Ecosystem is clear in the favour of London (MandagMorgen, 2015). Cultural differences have also been shown to have a strong impact of VC performance (Nahata, Sonali, & Kishore, 2014). Based on the 6 elements of Hofstede (2011), major differences in terms of culture has also been identified. Here, the main finding is that the United Kingdom is focused on competition and the desire to stand out from the crowd.

The **H3** was rejected as the difference in sample means were insignificant and therefore no evidence that Foreign experts were better than Danish experts. The insignificance is obviously a result of the small sample size of experts.

- The foreign crowd did in fact outperform the Danish crowd, with a total average return of almost 500 points.
- In addition, the findings presented a major difference in the amount of invested points in Company 1. Once again, the foreign experts were the best performing crowd.

Ecosystems in Silicon Valley, New York and London is some of the world's leading places to start and grow a company. The reason is partly because of the characteristics of the city, such as size, policies and location. Moreover, in regard to cultural differences, (Hofstede, 2011), there are clear differences between London and Copenhagen, which are the two Ecosystems presented in this study. Masculinity (Hofstede, 2011), in particular, is an element where London has a major advantage. London is focused around competitiveness and performance. A clear desire to be the best among others is a characteristic of UK culture. Therefore, it is argued that Denmark, should seek to foster a more competitive driven environment, that can enhance the performance of both VC's and start-ups.

Denmark, as one of the leading nations of numbers of start-ups produced, arguably has a large potential to enhance its global competitiveness by spotting the best start-ups. "*We're good at being innovative but not very good at scaling them"* Thomas Krogh Jensen said, CEO of Copenhagen Fintech (Keane, 2017). Copenhagen is argued to be a strong player in the early-stage scene but lack later stage funding (Keane, 2017). It is argued that for Copenhagen to act as an Ecosystem at a level equally to London, Copenhagen should foster an environment that are focused on competition and not being afraid from standing out from the Crowd. One way of dealing with this is to break with the "Jantelov" and foster an environment where it is acceptable to be different and in believe in your idea and identity. Additionally, Copenhagen should be better at keeping their resources in regard to investors and entrepreneurs, which could be done by creating better and more flexible conditions through enhancing stock market and legal protection (Nahata, Sonali, & Kishore, 2014), which have shown to have a positive impact for VC's performance (Nahata, Sonali, & Kishore, 2014).

Does experience make you a better forecaster?

H4 tested whether experienced experts were better predictors than less experienced experts. Experience within groups of experts is shown to have a positive effect on predictions (Mandel & Barnes, 2014). As venture capitalists evaluates differently depending on their experience (Franke et. al, 2008) and Sherpherd et al. (2002) showed a curvilinear correlation between experience it was investigated whether experienced had an effect on the predictions.

- The results showed a slightly better prediction, around 200 points, in favour of experienced experts. They perceived the two best companies almost equally, however, the worst company – Company 3, was invested the most in by less experience experts. As that specific start-up did not generate any return of investment, it is determined to be the reason for their marginally lower return on the total average investment. The results were insignificant.

The small sample size makes it hard for the researchers to discuss the findings. However, the topic of diversity within experts crowds are considered interesting to investigate. The experienced experts generated a marginally higher amount of points than less experienced experts. The difference in prediction was not more than 200 points, therefore, it is argued to be less important to have high degree of experience when judging start-ups. Sherpherd et al. (2002) has shown a curvilinear relationship between experience and predictions. Hence, it can be argued that Shepherd et al. (2003) is right about that there is a cap for maximum level of experience, as there was no significant difference in the predictions.

Summary

The relative small sample size makes it difficult for the authors to conclude anything major on **H3** and **H4**. To briefly sum up the discussion of **H3** and **H4,** internal resources of expert crowds were investigated. Experienced and foreign experts outperformed less experienced and Danish experts by a margin. The two hypotheses were interesting to the researchers, as the lack in performance of VC's, could be a result of their internal resources. The insignificant P value is a result of the low sample size. Despite the low sample size which leads to insignificant results, there are clear cultural differences between London and Copenhagen ecosystems in terms of competitiveness and desires in life (Hofstede, 2011). Copenhagen could potentially benefit from creating a stronger ecosystem by making it more attractive for high risk investors and entrepreneurs, which could also attract foreign companies to position themselves in Copenhagen. The sample size consisting of 3-4 experts from each expert crowd is the primary reason that the results showed insignificance. If an in-depth investigation of expert's crowds were to be done, the research would obviously need more expert participants. However, the small sample size highly reflects a real-life situation of a decision-making process in a VC. In the typical VC, there are usually two responsible individuals for each case.

Are experts better calibrated than non-experts?

H5 tested whether experts are better calibrated than non-experts. Experts are shown to be better calibrated than non-experts (Keren, 1987; Andersson et al., 2005; Whitecotton, 1996; Önkal et al., 2003). If experts judge on related items, it is shown that they can develop their ability as they develop experience (Keren, 1987). In this case, well calibrated mean that their confidence level reflects the amount of return of investment. E.g. with a confidence level of 7, they should generate somewhere between 6000-7000 points.

- Experts express a linear relation between their investment and their confidence level. Despite the fact that they were not able to outperform the total crowd nor the non-expert crowd, they are better calibrated. Although the results were insignificant, which primarily is caused by the relative low sample size, the correlation between the experts' confidence level and their investment were 0,66.
- There was no correlation between the confidence level of the non-experts and their investments. Approximately 50% of the non-experts express underconfidence.

These results could reflect the fact that the experts are generally more thoughtful in their decisions as they have prior experience within the industry. Venture Capitalist are experienced in developing probabilities of start-ups in terms of market fit and potential success of the team, as this is a core task of their job. Prior knowledge could therefore be argued to benefit the general task of creating probability estimations. However, it tends to result in overconfidence and therefore biased decisions. Hence, venture capitalists could benefit from training in debiasing (Tetlock, 2015).

Are predictions victim of confidence bias?

H6 tested whether experts and non-experts in general showed overconfidence in their predictions and whether experts were more overconfident than the non-experts. People are in general overconfident (Firschhoff, Slovic & Lichtenstein, 1977). Additionally, experts tends to express overconfidence in prediction tasks (Andersson, Edman & Ekman, 2005; Aukutsionek & Belianin, 2001; Braun & Yaniv, 1929; Kahneman & Riepe, 1988). Overconfidence especially is often seen in economic forecasts (Kahneman & Riepe, 1988). Overconfidence can occur due to e.g. overload of information and prior knowledge (Andersson et al., 2005).

- Findings showed that experts on average are overconfident as they generate less points than they indicated in their confidence level.
- 43% of experts was overconfident compared to only 29% of non-experts.
- Non-experts on average did not express overconfidence.

Despite the results being insignificant, the findings were in alignment with previous studies of experts and overconfidence (Kahneman & Riepe, 1988). Though, Fischhoff, Slovic and Lichtenstein (1977) argues that people in general are overconfident when they are certain to be right. As the non-experts did not show any clear pattern of a strong relationship between confidence and investment, the non-experts can be argued not to be certain about the outcome of the four start-ups. As expected the experts, determined to have prior knowledge within the field, showed overconfidence in their predictions. This might be a result of experts being to certain in their prior knowledge. As venture capitalists show overconfidence bias, venture capitals could arguably benefit from underestimating their venture capitalists' predictions, as "*the combination of overconfidence and optimism is a potent brew, which causes people to overestimate their knowledge, underestimate risks, and exaggerate their ability to control events"* (Kahneman & Riepe, 1988, p. 54). Additionally, training in debiasing would benefit VC's. Finally, VC's should hire people that are humble and open-minded, as Tetlock (2015) has proved individuals possessing these characteristics to be the best forecasters.

Summary

To sum up the discussion on **H5** and **H6**, the results were insignificant, however, they corresponded with previous literature. By dividing confidence level and total return into intervals, the experts in general expressed a linear correlation between confidence and investment. This is argued to be because experts possess prior knowledge of developing probability estimates, which they base their investments on. Though, they are generally overconfident. It could be argued that venture capitalists would actually benefit from being underconfident, as they invest in many companies that does not end up being successful. Hence, underconfidence could lead to venture capitals only investing in those companies they are 200% sure of. If venture capitalists were more conscious about this fact, this could create higher reflectiveness and lead them to active their slow thinking system2, instead of letting their fast thinking System1 getting the best of them (Kahneman, 2003).

Conversely, the non-experts on average showed a slight underconfidence. However, there was no linear relation between their confidence and investment, which mean they were not well calibrated.

Research have shown that too much information and knowledge can be misleading in predictions (Andersson et al., 2005). Non-experts does not have the same amount of information and knowledge to shuffle between and therefore does not have the possibility to compare with previous investments.

7.2 General Discussion

Most results of the thesis were found to be insignificant. However, the researchers argue that this is likely due to the sample size, which was a result of the timeframe for the thesis. As the results indicated that the law of large numbers had an effect on the predictability, the researchers argue that a larger sample size would result in significant results. The law of large numbers states that the mean should be closer to the mean of the whole population as the sample size increases (Routledge, 2018). When tested for the results showed a likewise development. By dividing the total sample into three equal groups depending on their timestamp of participation, the results showed that the collective accuracy increased in terms of spotting the right investments and the positive increase in significance level. Hence, the larger the crowd the more precise the crowd became in their predictability, as the average return of investment became better. This is consistent with the findings of Keuschnigg & Ganser (2017) that showed the larger volume of a crowd, the more accurate the crowd were in its predictions.

Team vs. Product

In order to investigate what underlying factors that drove the participants' investments, a multidimensional Likert-scale were developed. The Likert-scale were constructed in accordance to existing literature of venture capitalist's evaluation criteria and acted as an explanatory instrument for factors that correlated with the respectively investments. Hence, the Likert-scale were divided into "product" and "team". Results showed that a strong positive relation between "product" and the overall investment existed. Thus, "product" were the main indicator that drove the participants investments. This was found to be significant in all instances when tested on the total crowd. This was surprising to the researchers, as prior literature on venture capital and the interviews with investment managers all indicated that the team is the superior evaluation criteria compared to the product/idea. A recognized mantra within venture is the idea of rather investing in an A-team with a B-idea than in a B-team with an A-idea (Sherpherd & Zacharakis, 2001; Appendix 7). However, the circumstances of the thesis where cases and thereby also teams were presented on paper, seem to decrease the effect the start-up team have on the overall investments. This is understandable, as the information was gathered from LinkedIn, databases and company websites, which is likely only to reflect the experience and skills of the team and did not create a deeper understand on the personalities of the team. Additionally, the interviews of Danish investment managers revealed that "craziness" of the team members plays a large role when deciding which start-ups to invest in. In other words, "craziness" can be defined as the team members being special or different. This factor is hard to reflect in the presented cases as this factor relates to feelings, intuition and tacit knowledge. Yet, even as the "team" is seen to play a decreased role in guiding the investments, the results show that the crowd were still able to invest most points in the two "good" investments and least points in the "bad" investments. Therefore, a potential could lie in enhancing the "team" factor, as prior literature on venture capitals' evaluation criteria argue this to have a large effect on the decision process when judging start-ups. Thus, if participants have had access to a video or a face-to-face conversation with the team behind the four companies, it would possibly have an effect on the predictions. However, the correlation between the ratings of the "product" and the investments could also act as an alternative explanation for why the success rate of venture capital is relatively low. As the non-expert crowd were found to have a stronger correlation between the "product" and the investments, but still were found to be more accurate in their predictions, this could indicate that venture capitalist's places too great emphasize on the start-up's team as a reason to invest.

With this in mind, it is interesting to look into how the factor of the "product" could be enhanced. All four companies presented in the thesis were B2B (Business-to-Business) and can be argued to be harder for a layperson to understand, than if the cases had consisted of B2C (Business-to-Customer) companies. Companies with a B2C focus are often more relatable to laypersons, as these are concentrated directly at the customer. Hence, the customer has to have a clear understanding of the product in order to buy it, but do not necessarily have to be an expert in the underlying factors that drive the product's success. In addition, the final comments from the survey participants, which had the purpose of getting the opinions from the crowd members for the reason of their investment, showed that some participants found the products hard to understand (Appendix 8). This is understandable, as B2B products often can be directed at an underlying field that help drive the main product or be targeted at experts. Yet, the results showed that non-experts were more accurate in their predictions than experts. Thus, it is argued to be likely, that results would be even more accurate if the companies presented had been B2C as the non-experts could act as a potential end-customer.

Diversity

The theory of The Wisdom of Crowds argues that three factors have to be present, in order for the theory to be sufficient – diversity, dependency and decentralization

"*Diversity and independence are important because the best collective decisions are the product of disagreement and contest, not consensus or compromise. An intelligent group, especially when confronted with cognition problems, does not ask its members to modify their positions in order to let the group reach a decision everyone can be happy with*." (Surowiecki, 2004 p. 28)

Diversity is especially important in small groups (Surowiecki, 2004). Diversity is shown to be low in venture capital (Dober, 2017). Venture capital show a lack of diversity when it comes to race and gender, which is also reflected in the investment portfolio of venture capital (Truong, 2017). The diversity of the sample could definitely be greater; however, it is argued that the sample contains at least a higher degree of diversity than seen in the average VC. The sample consisted of both foreign and Danish participants, reflected a high degree of different backgrounds in regard to "field of expertise", had a wide range of age difference, though with a concentration of young people between the age of 19-38, and consisted of 1/3 females, which is a much higher concentration than the average of venture capitals. As the researchers looked into diversity they investigated whether any correlations could be found between demographics and the investments, however, the researchers found no strong correlations. This could indicate that diversity is the key to accurate predictions as certain key characteristics cannot explain the investment patterns. However, it could also be argued to be a result of the relative small sample size. Though, the results showed that females generated a higher average return of investment. This result was not significant and the excess amount they generated was not of high volume. Nonetheless, it could be a motivational factor for VC's to benefit from a higher degree of females within their investment teams. in general, it is argued that VC's should strive to increase diversity in investment teams.

Independence

Groups tends to come up with better results than individuals (Tetlock, 2015). However, finding the right balance within a group are not an easy task. If groups are insufficient in managing themselves in a proper way they can suffer from bias. Bias can occur e.g. from groupthink (Janis, 1972), that would lead groups not to exploit their cognitive diversity, eventually leading to incorrect predictions. Venture capitals act as small groups and with the low success rate in mind, it could be argued that groupthink, low diversity or other biases are presents, as they otherwise would be likely to produce better results. However, venture capitals could potentially benefit from holding their venture capitalists accountable of their predictions, while still obtaining independence, as accountability could lead to more precise predictions and probability developments (Tetlock & Kim, 1987). This study found that by having an independent crowd of non-experts they were able to produce results that found the best investment cases. If the participants were able to modify answers based on previous prediction, this would eliminate the crowd being independent. It is arguably very likely, that crowd participants would be biased by having the possibility to either modify their forecasts or cooperating in small groups and thereby possibly creating different results.

Decentralization

Decentralization states that the power of decisions should not be centralized around a few or even one individual. The results of this study could act as a decentralized decision making. In the industry of venture capital, decision makers might include advisors or colleagues to support the decision - Yet, these are often used as supplements to the final decision, which is taken by the CEO or few partners (Appendix 7). VC's could benefit from letting the solution being totally decentralized, either to their internal employees or to a external crowd, as they would avoid the final decision bias of a single decision maker. If a decentralized investment decisions was taken based on the findings from this study, investments would be made in Company 4 and alternatively, also Company 1. However, if an organization would base their decision only from collective wisdom, it could offend the traditional decision makers, as they could feel neglected. Therefore, VC's could benefit from restructuring the decision-making process to be more decentralized, by promoting the intelligence of large and diverse groups. In many other aspects, companies have been successful in this task, as e.g. crowdsourcing have become a more accepted way of obtaining new and good solutions over the last couple of years (Afuah & Tucci, 2012).Moreover, the term Crowdfunding has also gained an increased interest during the recent decade. Crowdfunding is a way to let the collective determine if there is a demand and market fit for a certain product by a large and diverse crowd. Though, Crowdfunding is not seen to be independent, as all possible investors are able to see the amount of capital raised for a certain start-up and can thereby be biased because of this.

A practical example of a VC that have increased the level of decentralization internally is Vækstfonden. When evaluating a start-up, all relevant investment managers and partners simultaneously turn a paper that show their particular "grade" of the start-up. By doing so, the investment managers and decision makers will not be biased by each other's opinions. This is, according to Vækstfonden, a relative new initiative, as they strive to improve internal processes, as they are in fact aware of internal bias (Appendix 7.2). However, venture capitalists are biased by the environment they act within. Hence, they are likely to invest in companies that they have heard other within the venture capital praise about. Availability bias argues that individuals will tend to judge situations from knowledge that are available and most current to them (Kahneman & Tversky, 1973). Hence, the conversation from e.g. the lunch break, could lead to biased judgments.

How to become a better forecaster

Findings of the study indicate that non-experts are more precise in their investments than experts. This is in strong opposition to the expert-view that constitutes the western society. The question is if experts are actually necessary if the crowd hold the more accurate knowledge on average. Tetlock (2005) found that experts within political judgment are not necessary as non-experts created better predictions than experts. However, he also found that individuals could become better at forecasting future political events based on their mindsets' pre-set when producing judgements. It matters less what experts think than how they think (Tetlock, 2015). Tetlock (2015) showed that when the goal is to produce accurate predictions, individuals who knows a lot about many things are better predictors compared to those who base their predictions on knowledge within one certain field or see everything through one lens – the last mentioned are often referred to as experts. Venture capitalists are likely to build their predictions on prior knowledge within the field of venture capital. This indicates that experts in venture capital might be unnecessary, unless they change their process of thinking. Additionally, it was shown that individuals could become better at predicting through a few actions. Simple training in improving probabilistic thinking and the understanding of removing cognitive biases can increase predictions with 10% (Tetlock, 2015). Additionally, forecasters can become better by keeping score of their judgments. Keeping a score can act as vital feedback, that can help forecasters become better over time. Hence, if forecasters are to revise their results and then predict again this can improve their results (Tetlock, 2015). Hence, it is argued that if the same non-expert crowd were exposed to a time series of predictions that they could revise and see the results of, they could become even more accurate in their predictions. Hence, in order to become an expert in forecasting the process of forecasting is more valuable than specific knowledge on the field.

Ranking

The scope of the project was to test whether certain crowds were better at predicting start-ups than others. Studies on The Wisdom of Crowds has also shown that crowd are able to generate accurate answers when ranking certain items (Lee et al., 2012; Grazing 2016). As the goal of this project was to see whether crowds were able to identify which start-ups were considered as the best investments, it would be relevant to ask them for the specific ranking. However, the authors found it more relevant, that the participants didn't have to create a specific ranking dictated by the authors but instead, had the possibility of ranking investments as equally good or bad. The findings showed that the total crowd and non-experts were capable of not only sorting away the "bad" investments but also choose to invest the most points on average in the two "best". In alignment with previous studies (Klein and Garcia, 2015), which have shown that non-experts are able to sort away bad ideas in a screening phase, the findings on this study were the same. In fact, the crowd also identified the two best start-ups.

Feedback

Moreover, different types of feedback have proven different effect of how increase the accuracy of forecast (Lawrence et al., 2006; Benson & Önkal, 1992). Outcome feedback is an example of a type of feedback (Benson & Önkal, 1992). Though it would be difficult to measure in a VC, as the usual fund has a lifespan of 10 years and additionally, start-ups usually generate its return of investment after at least 5 years (Appendix 7). Therefore, in a VC, it would be difficult to measure feedback only based on outcome, as industry trends change and because of the long time periods from the initial investment to the return of investment, through an exit or IPO, is realized. Many other types of feedback and ways to improve decision making in VC's has been studied. As outcome feedback would be too difficult to realize and the results would not be useful, VC's could break down the processes that leads to the decision and try to improve certain areas by measuring its performance. This could e.g. be the process of screening team members through several personality and/or intelligence test. Moreover, they could support their decisions with bootstrap models or AI methods as statistical model have proven to be very accurate in prediction task under high uncertainty.

Chapter 8 - Conclusion

The study set out to answer to the research question: '*How can crowds' predictions be used to optimize the low success rate of Venture Capital and are predictions victim of overconfidence?'*

In order to investigate whether crowd predictions can be used to optimize the low success rate of venture capital, the researchers conducted a quantitative crowd prediction tournament. The crowd prediction tournament explored both the use of an external crowd of non-experts and the effect of internal crowds of venture capital with different characteristics. The study established findings of non-expert crowd's ability to predict the success of start-ups compared to-, and in combination with experts. The findings indicated that a large, diverse and independent non-expert crowd are more accurate in predicting the success of start-ups than a small expert crowd and a combination of an expert-, and a non-expert crowd. This was found, as the non-expert crowd on average were able to generate a higher return of investment and were able to identify and invest the most on average in the two best start-ups out of the four start-ups presented in this case. The same findings were found reluctant of using the method of aggregating human judgments or the methods of calculating the probability of participants picking the best start-ups. However, the results were found to be insignificant, so no conclusions on this can be made. Yet, this was primarily an effect of the small sample size. Hence, if the sample size had been of greater volume the researchers would have been able to conclude on the effect crowd predictions have on venture capital's low success rate.

The study additionally, investigated if any internal characteristics within venture capital could work as a solution to the low success rate within venture capital. This was explored by dividing the experts into different groups depending on their relation to the characteristics explored, which was experience and nationality and then test the average prediction outcomes of generated return of investment. Even though the findings were insignificant, the results showed that foreign experts were able to generate a higher return than that of Danish venture capitalists. This indicates that foreign experts are likely to be better forecasters than their Danish colleagues. These findings were in alignment with the theory of Hofstede, who has identified cultural differences between Denmark and the United Kingdom that act in favour for London as a strong Ecosystem.
The findings of experience showed that experts holding at least 10 years of experience generated a slightly higher return of investment than their less experienced colleagues. However, the extra amount generated on average was off so little margin, that the results indicate that experience have no effect.

Both results were insignificant because of the small sample size of experts, which was an effect of the scope of the thesis. Nonetheless, the small sample size is a good reflection of reality, where only few venture capitalists are to perform the evaluation- and investment processes within VC's. Still, the researchers acknowledge that in order to gain a deeper understanding of whether internal characteristics within venture capital have an effect on their predictions a greater volume of experts is needed to create sufficient results.

To investigate whether crowd predictions are victims of overconfidence, the researchers explored the accuracy of the participants predictions compared to their subjective perception of the prediction. The results indicated that the experts on average suffered from overconfidence. However, it was also found that experts were better calibrated than non-experts. In fact, findings showed that non-experts on average showed an underconfidence in their predictions. However, none of these results were found to be significant. Though, the results showing experts to be overconfident corresponds well with low success rate within the industry of venture capital. This is argued as VC's are unlikely to invest in start-ups they do not believe in but with so many venture backed start-ups failing to generate the expected return, this must be a result of overconfidence. Hence, it is argued that the low success rate of venture capital could be enhance from venture capitalists being training in debiasing their overconfidence.

Finally, the researchers explored what underlying evaluation criteria that drove the investment predictions, here *the product* and *the team* was investigated. The researchers found significant results of *the product* being the evaluation criteria that mainly drove the investments. Additionally, it was found that experts placed a slightly greater emphasize on *the team* than the non-experts and as non-experts were able on average to create more accurate predictions than those of experts, it is argued that VC's can benefit from placing more emphasis on *the product* when evaluating start-ups. However, the researchers acknowledge the limitations of the presentation of the respective start-ups and are not able to conclude specific on *the product's* effect on venture capitals low success rate. Nonetheless, as the evaluation criteria of venture capitalists haven't yet been explored in relation to the success these create, it could be argued that a further investigation of which evaluation criteria that actually are used when start-ups turn out to be successful and thereby hold a high potential for eventually improving the success rate of VC's.

As a final remark, the results indicate that *The Wisdom of Crowds* holds a high potential to increase the low success rate of venture capital. However, as the results from this study were found to be insignificant and all of the six hypotheses were rejected, the researchers encourage other researchers with a larger time-frame of their project to test the effect of *The Wisdom of Crowds* within venture capital on a larger scale.

Chapter 9 - Limitations

One of the three main characteristics of Wisdom of the Crowd is diversity. The sample is argued to hold a sufficient degree of diversity. However, the degree of diversity could definitely be increased. Due to the scope of the project, the researchers were forced to use e.g. public pages on Facebook to get participants for the survey. This has resulted in a cluster in the age difference reflected in the study, as many participants are between the age 19-38. The researchers tried to overcome this age cluster by acquiring participants from "Strøget", as this place was believed to reflect many different age differences. Though, it should be mentioned that the researchers are not able to track if the participants are acquired from "Strøget", public Facebook pages or one of the other sources used to gain participants. However, this only goes for the non-expert crowd. The experts were found through email correspondence and the researcher are able to trace back to them. Additionally, the researchers got four tutors to spread the questionnaire on their study's internal Facebook pages, which could again limit the diversity of the crowd, as most participants from these groups are likely to be from Denmark and have the same educational background. However, due to the scope of the thesis the researchers had to use their networks in a close proximity to them, which accounts for the geographical cluster of participants.

Another limitation to the study is the volume of experts, as only 7 experts participated in the survey. The number of participants decreases the validity of the study, as these experts do not necessarily reflect venture capitalists comprehensively. Additionally, it decreased the usefulness of hypothesis 4 and 5, where the objective was to test characteristics of venture capitalists that could affect their predictions. Nonetheless, does the volume of the expert crowd reflect the reality rather good. Venture capital are often small organization that have investment team's that consist of a small group of people (Appendix 7.1). Hence, the results reflected in the thesis are likely to reflect an investment team within a venture capital. The expert crowd consisted of white males of different age, with different levels of experience and management involvement. This is the most common composition of venture capital teams. Therefore, venture capitals are extremely dependent on the members of their investment team. If just a few members demonstrate poor decision making the whole venture capital are likely to be affected by the decision of these. Generally, this is the case of many fields where experts are considered the solution. If few experts are present within a certain field, they are in danger of generating poor decision making, due to bias or the few individuals effect on others' judgments. Off course it is possible to find examples that testifies for the strength of experts. An example is

Warren Buffet, who by many people is believed to be the greatest stock investor of all times and over time again and again have outperformed the market. However, the truth is that very few examples such as Warren Buffet exists. The researchers are not at all neglect to the fact, that some venture capitalists might act as Warren Buffet match within venture capital. Though, very few are likely to do predictions at this level. Hence, it is argued that venture capitals are better off exercising the law of large numbers.

The researchers found that the strongest correlation to drive the participants investments were the “product”. This indicates that the value of the “team” have been decreased through the thesis presentation of cases. This could act as a limitation to the study, as the “team” generally in literature and through the conducted interviews is an extremely valued evaluation criteria to judge investment decisions by. Thus, it is likely that the predictions could have become more accurate, if the assessment of the “team” had been valued higher in the study. The researchers believe that the quality of the thesis would have increased if the team could have been presented in a video, so the participants could get a deeper understanding of the team behind the startup. Or if the founders had contributed with additional knowledge of themselves. But, this information was simply not identifiable or present to the researchers. The researchers explored many possibilities to develop the cases presented in the thesis. Nonetheless, venture capitals have very strict and clear confidentially agreements, that cannot be waved even though the material needed was historically. But, if the researchers were to do a larger study investigating the same effects, it would be a priority to use material that could enhance the understanding of the team behind the start-up. It is argued that this could enhance the quality of the results.

As the cases presented are historical, this represent a limitation. The historical perspective of the cases could act as an effect on the participants predictions. Participants have been able to observe what trends and business areas that have developed the most over the period of time, since the companies got their initial investments in 2011, which could have an effect on their ability to predict what companies have become success’. Nonetheless, the companies all act within business areas that are still very potent and hold a high potential and it is argued it would be difficult to identify the best company based only on this. The researchers actively picked companies within business areas that still holds a high potential in order to adjust for this limitation. Another limitation posed by the historical perspective is the possibility of individuals being able to identify the specific companies.

Therefore, the identity of the companies, the founders and the venture capital that had invested in them was excluded from the cases. Additionally, participants were asked not to invest if they had prior knowledge of the specific company. Hence, this limitation is argued to be adjusted for and eliminated to the researcher's best ability.

Finally, the historical perspective could pose a limitation to the validity of the information gathered and presented about the companies. The information might have varied to a certain degree since the initial investment was given in 2011. However, the researchers found that the information in the cases correspond rather well with the information gained from Almi Invest. Additionally, the researcher's included only information about the founders from the period before and until 2011 and tried to their best ability to exclude information that was about the period after the initial investment.

The researcher's initial idea was to obtain predictions based on companies that were in the early startup phase in present time, as this would exclude the historical perspective of the cases and probably enhance the availability of data and information. However, as success in regard to venture capitals 10x rule can only be measure after a longer time period this idea was excluded as no results would be found. The researcher thought of correcting for this by looking at four start-ups where some of them had just obtained a series A funding and other had been denied it. However, this would not indicate that the start-ups would reach success in regards of the 10x rule in the long run. Hence, the pros and cons of the historical limitation was considered intensely and found to be the best solution for testing Wisdom of the Crowds effect on venture capital.

Chapter 10 - Future studies

The scope of the thesis only allowed the researcher to measure on participant's prediction at one point in time. A series of predictions would have been unrealistic, as it is not easy to get participants to answer surveys. The researchers had to post the survey on a regular basis and over a quite long-time period in order to obtain the relative limited responses of the study. Hence, it would have been unrealistic to have measured more than once within this scope of the thesis. However, it is argued that measuring over a time series could enhance the quality of the results. Measurement over a time series could take different shapes. It would be interesting for future research to test the same cases on more than one sample to test whether the same results would occur and enhance the data quality of the study. Additionally, it would be interesting to use other case examples to test whether a crowd could come up with the same accuracy in their predictions. Additionally, the researchers would suggest that cases were performed in a way that could enhance the presentation of the team behind the start-ups. This could be done through video introduction, letter presentation by the founder/founders or if future researchers could gain access to original material in terms of one pagers, PowerPoint shows etc., this would definitely give a better reflection of the team behind and could affect the results.

Furthermore, it would be interesting to do a research project, such as a Ph.D. that runs over three years, to test whether participants can predict what companies are likely to be most successful over a time period of three years. This would create the most reliable results, as information sources such as trends in the society etc. could be up to date and taken into consideration.

The above mentioned are all examples of how future forecasting studies could be structured when no markets are present, as the case of this study. However, the researchers would also encourage future studies to test the effect of The Wisdom of Crowds on venture capital when forecasting markets are present. This could be done through the use of prediction markets. A study of this kind, could test prediction markets effect compared to no markets, if an initial independent prediction was done and then compared to the results of the prediction market. A study of this kind would require a large timeframe as the prediction market would have to run over a longer period of time, in order to indicate significant findings.

This study had it focus on human capabilities of forecasting within venture capital. Yet, the researchers acknowledge the high potential of AI. Hence, it could be interesting to create a future study that combines the effects of The Wisdom of Crowds and the abilities of AI. The combination would be likely to create the most interesting results for future venture capitals.

Finally, the researchers would encourage future researchers to focus only on the characteristics of venture capitalists that enhances their performance as forecasters, in order to investigate if any of these are present. Hence, in order to create significant results of the tested hypothesis 3 and 4 from this study, researchers would have to create a larger experts crowd. This could help future venture capitals to understand if they could hire venture capitalists based on certain characteristic that could enhance the venture capitals performance. E.g. if experience, background or other factors were to play a role.

Chapter 11 - Practical implications

Findings indicated that non-experts were more accurate in their prediction of start-ups success, however, the results insignificant. If future studies could prove significance venture capitals should consider hiring a non-expert crowd to create predictions on their cases. The crowd could act totally as venture capitalists do today or could be used as a supplement to the expert's assessments. Alternatively, a non-expert crowd could act as the first line of soldiers, that filters away the bad ideas, as crowds are better at rejecting bad ideas, than they are finding the best ideas (Klein & Garcia, 2015).

Secondly, the research implied that experts suffers from overconfidence. This corresponds with earlier findings. Tetlock (2015) found that predictions could become better by the use of simple training in bias removal. Even though the results didn't show significant overconfident bias, the researchers would argue that the low success rate within venture capital is a logical argument for overconfidence, as venture capitalists wouldn't be likely to invest in companies that they are not confident will succeed. Thus, it is argued that venture capitals could benefit from training their venture capitalists in how to reduce biases.

Finally, the results showed that the "product" were the main indicator to drive an investment and not the "team". This could indicate that venture capitalists should enhance their focus on the initial idea, and not place to great emphasize on the "team" behind the start-up. Nevertheless, the researchers know that the presentation of the "team" lacked in information and are not encouraging to neglect the value of the "team".

References

(u.d.). Hentet fra https://www.mckinsey.com/industries/high-tech/our-insights/a-machine-learning-approach-to-venture-capital

Afuah, A., & Tucci, C. (2012). CROWDSOURCING AS A SOLUTION TO DISTANT SEARCH.

Alper, S., Tjosvold, D., & Law, K. (1998). Interdependence and Controversy in Group Decision Making: Antecedent to Effective Self-Managing Teams. *Organization Behaviour and Human Decision Processes*.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior Volume 22*.

Andersson, P., Edman, J., & Ekman, M. (2005). *Predicting the World Cup 2002 in soccer: Performance and confidence of experts and non-experts*.

Armstrong, S. (1983). Relative accuracy of judgemental and extrapolative methods in forecasting annual earnings. *Journal of Forecasting*.

Armstrong, S. (2001). *Principles of Forecasting: a handbook for researhers and practitioners.* Springer Science + Business Media.

Aukutsionek, S., & Belianin, A. (2001). Quality of forecasts and business performance: A survey study of Russian managers. *Journal of Economic Psychology Volume 22, Issue 5*.

Axelrod, R., & Douglas, D. (1988). The Further Evolution of Cooperation. *Science* .

Önkal, D., & Muradoglu, G. (1994). Evaluating Probabilitstic forecast of stock prices in a developing stock market . *European Journal of Operational Research*.

Önkal, D., Yates, F., Simga-Mugan, C., & Oztin. (2003). Professional vs. amateur judgment accuracy: The case of foreign exchange rates. *Organizational Behavior and Human Decision Processes 91* .

Bachher, J., & Guild, P. (1996). FINANCING EARLY STAGE TECHNOLOGY BASED COMPANIES: INVESTMENT CRITERIA USED BY INVESTORS. *Frontiers of Entrepreneurship Research*.

Bendsen, T. (2018). *statnoter.dk*. Hentet fra statnoter: www.statnoter.dk

Benson, G., & Önkal, D. (1992). The effect of feedback and trianing on the performance and probability forecast. *International Journal of forecasting 8*.

Berstein, S., Korteweg, A., & Laws, K. (2016). Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment. *The Journal of Finance*.

Blohm, I. (2016). Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing. *Institute for Operations Research and the Management Sciences*.

Bonabeu, E., & Meyer, C. (2001). *Swarm Intelligence: A whole new way to think about business*.

Bon, G. L. (1995). *The Crowd.* New York.

Braun, P., & Yaniv, I. (1992). A case study of expert judgment: Economists' probabilities versus base-rate model forecasts. *Journal of Behavioral Decision Making* .

Brenner, L., Griffin, D., & Koehler, D. (2005). Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgment. *Organizational Behavior and Human Decision Processes, Vol, 97, Issue 1*.

Brunsson, N. (1982). Journal of Management Studies. *THE IRRATIONALITY OF ACTION AND ACTION RATIONALITY: DECISIONS, IDEOLOGIES AND ORGANIZATIONAL ACTIONS*.

Bruton, G., & Ahlstrom, D. (2003). An institutional view of China's venture capital industry: Explaining the differences between China and the West. *Elsevier*.

Bruton, G., & Rubanik, Y. (2002). Resources of the firm, Russian high-technology startups, and firm growth. *Journal of Business Venturing 17*.

Bryman, A., & Bell, E. (2003). *Business Research Methods.* Oxford University Press.
Chakrabarti, R., Gupta-Mukherjee, S., & Jayaraman, N. (February 2009). Mars–Venus marriages: Culture and cross-border M&A. *Journal of International Business Studies*.
Chin, W. (1998). The partial least squares approach to structural equation modeling. *Modern methods for business research*.
Ciszewski, B. (2016). *Why London Startup Ecosystem Might Be The Best In The World*. Hentet fra Netguru: https://www.netguru.co/blog/why-london-startup-ecosystem-might-be-the-best-in-the-world
Clark, B. (2016). *Evaluating a Startup? These 7 Metrics Can Help*. Hentet fra Micro Ventures: https://microventures.com/evaluating-a-startup-these-7-metrics-can-help
Cooper, A., & Bruno, A. (1977). Success Among High-Technology Firms. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.
CreativeResearchSystems. (2018). *surveysystem*. Hentet fra surveysystem.com: https://www.surveysystem.com/correlation.htm
Csaszar, F., Nussbaum, M., & Sepulveda, M. (2006). Strategic and cognitive criteria for the selection of startups. *Elsevier*.
Dapresy. (2018). Hentet fra Dapresy: https://www.dapresy.com/
Dawes, R. (1974). Linear models in decision making. *Psychological Bulletin, 81(2)*.
Dealroom. (2017). *Fundraising by European venture capital funds.* Dealroom.co.
Dober, S. O. (21. September 2017). *Diversity is lacking in venture capital. It's time to change that. | Detroit VC leaders* . Hentet fra Freep: https://www.freep.com/story/opinion/contributors/2017/09/21/venture-capital-diversity/690328001/
Dreyfus, H., & Dreyfus, S. (1986). *Mind over Machine.* New York: The Power of Human Intuition and Expertise in the Era of the Computer.
Edmundson, H. R. (1990). Decomposition; a strategy for judgemental forecasting. *Journal of Forecasting, Vol 9*.
Farha, D. (29. May 2016). *Understanding The VC Business Model*. Hentet fra Entrepreneur: https://www.entrepreneur.com/article/276639
Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with Certainty: The Appropriateness of Extreme Confidence. *Journal of Experimental Psychology: Human Perception and Performance, Vol. 3*.
Forrest, D., & Simmons, R. (2000). *Forecasting sport: the behaviour and performance of football tipsters*.
Franke, N., & Gruber, M. (16. April 2008). *Venture Capitalists' Evaluations of Start-Up Teams: Trade-Offs, Knock-Out Criteria, and the Impact of VC Experience*.
Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (16. April 2008). Venture Capitalists' Evaluations of Start-Up Teams: Trade-Offs, Knock-Out Criteria, and the Impact of VC Experience. *ENTREPRENEURSHIP THEORY and PRACTICE*.
Fredriksen, O. K. (1989). Growth, performance and financial structure of new technology based firms: A preliminary analysis. *Frontiers of entrepreneurship research, pp. 189-199*.
Gage, D. (19. September 2012). *The Venture Capital Secret: 3 Out of 4 Start-Ups Fail*. Hentet fra The Wall Street Journal: https://www.immagic.com/eLibrary/ARCHIVES/GENERAL/GENPRESS/W120919G.pdf
Galton, F. (1907). *Vox Populi* .
Generozova, O. (u.d.).

Gigerenzer, G., & Selten, R. (2002). *Rethinking Rationality.* Berlin, Germany: Center for Adaptive Behavoir and Cognition .
Glaser, M., Langer, T., & Martin, W. (12. October 2012). *True Overconfidence in Interval Estimates: Evidence Based on a New Measure of Miscalibration*.
Goodman-Delahunty. (2010). Insightful or Wishful: Lawyers' Ability to Predict Case Outcomes. *Psychology, Public Policy, and Law, Vol. 16, No. 2*.
Goodwin, P., & Wright, G. (1993). Improving judgmental time series forecasting: A review of the guidance provided by research . *Department of Mathematical Sciences,* .
Gordon, K. (1924). Group Judgments in the Field of Lifted Weights. *Journal of Experimental Psychology, 7(5), 398-400*.
Graham, M. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology*.
Griffith, E. (25. September 2014). *Why startups fail, according to their founders*. Hentet fra Fortune: http://fortune.com/2014/09/25/why-startups-fail-according-to-their-founders/
Guba, E., & Lincoln, Y. (1994). *Competing Paradigms in qualitative research.*
Guo, S. (2015). *Not Every Startup Should Be Venture Backed*. Hentet fra Medium: https://medium.com/@saranormous/not-every-startup-should-be-venture-backed-970d728059d
Hallin, C., Øgaard, T., & Marnburg, E. (2009). Exploring Qualitative Differences in Knowledge Sources: A Study of Hierarchical Effects of Judgmental Confidence and Accuracy Performance. *International Journal of Knowledge Management*.
Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations.* Sage Publications .
Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture, 2*.
hofstede-insights. (2018). *www.hofstede-insights.com*. Hentet fra www.hofstede-insights.com: www.hofstede-insights.com
Hogarth, R. M., & Makridakis, S. (1981). Forecasting and planning: An evaluation. *Management Science Vol. 27*.
Hong, H., Wang, A., Xu, D., Du, Q., & Fan, W. (2016). Crowd Wisdom: The Impact of Opinion Diversity and Participant Independence on Crowd Performance. *Twenty-second Americas Conference on Information Systems*.
Hong, L., & Page, S. (March 2001). Problem Solving by Heterogeneous Agents. *Journal of Economic Theory* .
Hong, L., & Page, S. (2004). Groups of Diverse Problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*.
Howe, J. (2006). The Rise of Crowdsourcing. *Wired Magazine - Issue 14.06*.
Hurree. (u.d.). *5 Current Limitations of Artificial Intelligence to Marketers*. Hentet fra hurree: http://blog.hurree.co/blog/limitations-of-artificial-intelligence-mobile-marketing
Huymans, J. H. (1970). The Effectiveness of the cognitive style constraint in implementing operations research proposals. *Management Science*.
Imamuddin, K. (2009). PAKISTANI VENTURE CAPITALISTS' INVESTMENT CRITERIA: A COMPARATIVE LOOK . *Asian Academy of Management Journal, Vol. 14, No. 1,*.
Intrade. (2018). *Intrade*. Hentet fra http://www.intrade.com/
Janis. (1972). Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes. *PsycINFO*.
Jones, B. D. (1999). Bounded Rationality. *Department of Political Science*.
Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. . *American Psychologist, Vol 58(9)*.

Kahneman, D., & Riepe, M. (1998). Aspects of investor psychology. *Journal of Portfolio Management*.
Kahneman, D., & Tversky, A. (1973). *Judgment under uncertainty: Heuristics and Biases.*
Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80(4), 237-251*.
Keane, J. (2017). *Copenhagen's startup ecosystem finds that its time to scale and mature*. Hentet fra techeu: tech.eu
Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study . *Organizational Behavior and Human Decision Processes*.
Keuschnigg, M., & Ganser, C. (2017). *Crowd Wisdom Relies on Agents' Ability in Small Groups with a Voting Aggregation Rule*.
Klein, M., & Garcia, A. (2015). High-speed idea filtering with the bag of lemons. *Decision Support Systems vol 78*.
KPMG. (2017). Handbook to Nordic startup investor market: unleash the full potential of Nordic startups. *Nordic Innovation 2017*.
Landemore, H. (2008). Democratic Reason: the Mechanisms of Collective Intelligence in Politics. *Department of Government Harvard University*.
Larrick, R., & Soll, J. (2006). Intuitions About Combining Opinions: Misappreciation of the Averaging Principle. *Management Science 52(1):111-127*.
Latsis, S. (1981). *Method and Appraisal in Economics.* JSTOR.
Lawrence, M., Paul, G., O'Connor, M., & Önkal, D. (2006). *Judgmental forecasting: A review of progress over the last 25 years*.
Lee, M., Steyvers, M., de Young, M., & Miller, B. (2012). Inferring Expertise in Knowledge and Prediction Ranking Tasks. *Department of Cognitive Sciences*.
Lemkin, J. (23. October 2012). *The 10x Rule: What Raising $1 of Venture Capital Really Means*. Hentet fra SaaStr: https://www.saastr.com/the-10x-rule-what-raising-1-of-venture-capital-really-means/
Lévy, P. (1997). *Collective Intelligence*.
Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know. *Oprganizational Behavior and Human Performance*.
Lichtenstein, S., Fischhoff, B., & Phillips, L. (1976). *Calibration of Probabilities: The State of the Art.*
Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry.* Sage Publications.
Lyon, A., & Pacuit, E. (2013). The Wisdom of Crowds: Methods of Human Judgement Aggregation. *Philosophy*.
MacMillan, I. C., Siegel, R., & Narasimha, S. (1985). CRITERIA USED BY VENTURE CAPITALISTS TO EVALUATE NEW VENTURE PROPOSALS. *Journal of Business venturing*.
Mahajan, V., & Wind, J. (1992). New product models: Practice, shortcomings and desired improvements. *Elsevier*.
Malone, T. W. (16. April 2009). *Harnessing Crowds: Mapping the Genome of Collective Intelligence*.
MandagMorgen. (2015). *Vækst-Danmark 2015.* Mandag Morgen.
Mandel, D. R., & Barnes, A. (2014). *Accuracy of forecasts in strategic intelligence*.
Mannes, A. E., Soll, J., & Larrick, R. (2014). *The Wisdom of Select Crowds*.
M. J. Lawrence, R. H. (1986). The Accuracy of Combining Judgemental and Statistical Forecasts. *Management Science Vol. 32, .*
McGrath, J. E. (1984). *Groups: Interaction and Performance.*

Meisler, L. (25. May 2016). *Who Gets Venture Capital Funding?* Hentet fra Bloomberg: https://www.bloomberg.com/graphics/2016-who-gets-vc-funding/

Miller, C., & Ireland, D. (2005). Intuition in strategic decision making: Friend or foe in the fast-paced 21st century? *Academy of Management*.

More, L. (2017). *The Hill*. Hentet fra http://thehill.com/blogs/pundits-blog/technology/327212-supporting-the-startup-economy-can-create-1-million-new

Mullan, E. (November 2017). *5 Current Limitations of Artificial Intelligence to Marketers*. Hentet fra Hurree: http://blog.hurree.co/blog/limitations-of-artificial-intelligence-mobile-marketing

Murphy, A., & Winkler, R. (1983). Probability Forecasting in Meteorology. *Journal of the American Statistical Association Volume 79,*.

Muzyka, D., Birley, S., & Leleux, B. (1996). Trade-offs in the investment decisons of European venture capitalists. *Journal of Business Venturing, Vol 11*.

Nahata, R., Sonali, H., & Kishore, T. (2014). Success in Global Venture Capital Investing: Do Institutional and Cultural Differences Matter? *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 49*.

Nelson, M., Libby, R., & Bonner, S. (1995). Knowledge Structure and the Estimation of Conditional Probabilities in Audit Planning. *THE ACCOUNTING REVIEW Vol. 70, No. 1*.

Neo4j. (2018). *Neo4j*. Hentet fra https://neo4j.com/

Nisen, M. (2014). *Bill Gates says the success rate on venture capital is "pathetic" compared to development*. Hentet fra Qvartz: https://qz.com/187959/bill-gates-says-the-success-rate-on-venture-capital-is-pathetic-compared-to-development/

O'Hear, S. (2017). *Fly Ventures, a Berlin-based VC using machine learning to find its next deal, closes 41M fund*. Hentet fra TechCrunch: https://techcrunch.com/2017/12/21/fly-ventures/

outsideinsight. (u.d.). *The growing focus on Artificial Intelligence in Venture Capital*. Hentet fra outsideinsight: https://outsideinsight.com/insights/outside-insight-in-venture-capital/

Palmer, M. (2017). *Artificial intelligence is guiding venture capital to start-ups*. Hentet fra Financial Times: https://www.ft.com/content/dd7fa798-bfcd-11e7-823b-ed31693349d3

Petersen, M. A. (July 2004). *Information: Hard and Soft*.

Pitchbook. (2018). *Pitchbook*. Hentet fra Pitchbook: www.pitchbook.com

Polanyi, M. (1966). *The Tacit Dimension.* Chicaco.

Quarterly, M. (u.d.). *A machine-learning approach to venture capital*. Hentet fra mckinsey: https://www.mckinsey.com/industries/high-tech/our-insights/a-machine-learning-approach-to-venture-capital

Rajan, T. (2010). Venture Capital and efficiency of portfolio companies. *IIMB management review*.

Rea, R. H. (1989). Factors affecting success and failure of seed capital/start-up negotiations. *Journal of business Venturing*.

Reddy, S. (2017). *A computer was asked to predict which start-ups would be successful. The results were astonishing*. Hentet fra World Economic Forum: https://www.weforum.org/agenda/2017/07/computer-ai-machine-learning-predict-the-success-of-startups/

REddy, S. R. (u.d.). *A computer was asked to predict which start-ups would be successful. The results were astonishing*. Hentet fra World Economic Forum: https://www.weforum.org/agenda/2017/07/computer-ai-machine-learning-predict-the-success-of-startups/

Routledge, R. (2018). *Central Limit Theorem*. Hentet fra Encyclopædia Britannica: www.britannica.com

Shaw, M. E. (1932). A Comparison of Individuals and Small Groups in the Rational Solution of Complex Problems. *The American Journal of Psychology, Vol. 44, No. 3*.

Shepherd, D., & Zacharakis, A. (2002). Venture capitalists' expertise A call for research into decision aids and cognitive feedback. *Journal of Business Venturing 17*.
Shepherd, D., Zacharakis, A., & Baron, R. (2003). VCs' decision processes: Evidence suggesting more experience may not always be better.
Shirley, D. A., & Langan-Fox, J. (1. October 1996). *Intuition: A Review of the Literature*.
Siegel, R., Siegel, E., & MacMillan, I. (1993). Characteristics distinguishing high-growth ventures. *Journal of Business Venturing Vol 8* .
Simmons, B., Lanuza, D., Fonteyn, M., Hicks, F., & Holm, K. (2003). Clinical Reasoning in Experienced Nurses. *Western Journal of Nursing Reseach*.
Simon, H. A. (1987). *Making Management Decisions: The Role of Intuition and Emotion*.
Song, M., Podoynitsyna, K., van der Bij, H., & Halman, J. (2008). Success Factors in New Ventures: A Meta-analysis. *Product Development & Management Association*.
StartupGenome. (2017). *Global Startup Ecosystem Report 2017.* Startup Genome.
Steiner, I. D. (1972). *Group process and productivity.* New York: Acedemic Press.
Stone, E., & Opel, R. (2000). Training to improve calibration and discrimination: the effects of performance and environmental feedback. *Organisational behaviour and human decisions processes*.
Sundali, J., & Atkins, A. (1994). *Expertise in Investment Analysis: Fact or Fiction*.
Surowiecki, J. (2004). United States: Anchor Books.
Surowiecki, J. (2004). p. 86. *Wisdom of Crowds*.
Telcred. (2018). Hentet fra www.telcred.com
Tetlock, P. E. (2005). *Expert Political Judgment How Good Is It? How Can We Know?*
Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The Art and Science of Prediction.* New York.
Truong, K. (4. January 2017). *Lack of VC diversity: New data shows that, yes, it's still a problem* . Hentet fra Bizjournals: https://www.bizjournals.com/sanfrancisco/news/2017/01/04/vc-venture-capital-diversity-tech-workforce-sf.html
Tyebjee, T. T. (1984). INFORMS.
Tyebjee, T., & Bruno, A. (1984). A Model of Venture Capitalist Investment Activity. *Management Science, Vol 30*.
Vertinsky, P., Kanetkar, V., Vertinsky, I., & Wilson, G. (1986). Prediction of wins and losses in a series of field hockey games: A study of probability assessment quality and cognitive information-processing models of players. *Organizational Behavior and Human Decision Processes Volume 38, Issue 3,*.
Waldron, J. (1. November 1995). *The Wisdom of the Multitude*.
Wauters, R. (2016). *TechEU*. Hentet fra http://tech.eu/brief/european-startup-monitor-2015/
Wells, W. A. (1974). Venture capital decision making.
Wheeler, W. M. (1911). *The ant-colony as an organism.*
Whitecotton, S. M. (1996). The Effects of Experience and a Decision Aid on the Slope, Scatter, and Bias of Earnings Forecasts. *Organizational Behavior and Human Decision Processes Volume 66, Issue 1,*.
Woolley, A. W. (2010). *s Evidence for a Collective Intelligence Factor in the Performance of Human Groups*.
Zacharakis, A., & Meyer, D. (2000). *The potential of actuarial decision models: Can they improve the venture capital investment decision?*
Zacharakis, A., & Sherpherd, D. (2001). THE NATURE OF INFORMATION AND OVERCONFIDENCE ON VENTURE CAPITALISTS' DECISION MAKING. *Elsevier*.

Zwilling, M. (25. December 2013). *A New Era For Entrepreneurs And Startups Has Begun*. Hentet fra Forbes: https://www.forbes.com/sites/martinzwilling/2013/12/25/a-new-era-for-entrepreneurs-and-startups-has-begun/#6c7f0514bd18