MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Studying Search in Structures - On the Effects of Accountability"

verfasst von / submitted by

Robert Janjic, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 915
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Betriebswirtschaft
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Chem. Dr. Markus Georg Reitzig, MBR

Abstract

Whenever people face a problem or have a task to perform, the need for finding a good solution arises. As numerous problems can be solved in various different ways, many of which potentially remain undiscovered for long periods of time, identifying the best solution is rarely a trivial task. The number of possible different solutions typically vastly exceeds the cognitive abilities of humans and frequently even machines. Constraints on resources such as time and processing capacities make an exhausting assessment of every thinkable solution to a problem practically impossible. In such situations, it becomes apparent why search is such a fundamental process in human decision-making. The discovery and evaluation of alternatives is a central aspect in countless activities ranging from simple everyday tasks to contexts of organizations taking decisions that affect a large number of people and can have far-reaching consequences. In the light of these examples, it becomes clear that studying human search behavior bears the potential for impactful discoveries and critical insights that can benefit decision-makers in a multitude of situations. Building on existing academic work in the realm of search behavior, this study aims at expanding current knowledge in the field by experimentally investigating differences in search strategies between individuals, unstructured groups and structured collectives.

Table of Contents

1 Introduction

Search is a fundamental process in human decision making as problems and their associated solution spaces usually by far exceed the processing capacities and resource constraints of human agents, giving rise to the need of screening a limited number of alternative solutions in order to find an appropriate way to solving a task or problem (Billinger, Stieglitz & Schumacher, 2014). If then, additionally, the solution space becomes complex due to interdependencies between decision parameters involved, the challenge of identifying effective strategies becomes even more demanding. Such interdependencies complicate relations between certain aspects of an existing solution and the associated contribution to the overall performance of this solution. When assigning the success of a strategy to its single constituent parts becomes non-trivial, then, anticipating the effects of changes in the strategy becomes exceedingly difficult. Suddenly, a small adjustment in one decision variable or a change in the configuration of the current solution to a problem can cause major distortions in the performance associated to this strategy (Levinthal, 1997).

A useful way to think about problems of that type and the related process of search for attractive solutions is the idea of a fitness landscape. Every possible solution to a problem is represented by a point on a landscape and the height of this point corresponds to its fitness or performance. Agents now "walk" over this landscape looking for promising locations, which represent a certain strategy or way of solving the task at hand (Levinthal, 1997; Kauffman & Weinberger, 1989). A possibility of "constructing" such a landscape is provided by the NK-model, introduced by evolutionary biologist Stuart Kauffman (1993). In this model, every decision that has to be made in regard to solving a problem or task is formulated in terms of a binary variable, i.e. a variable that can only take two values. In a practical, business-related context this could be a decision to implement a certain technology or not, to add a new product to the product line or not or other similarly structured decisions. All possible combinations of such decision variables involved are then mapped onto a landscape with their corresponding payoff or result they produce. Interactions between individual choices determine how "rugged" the landscape then is and, accordingly, how difficult it is to find peaks, i.e. high-performing solutions (Baumann, Schmidt & Stieglitz, 2019).

The important that then arises in the context of complex problem solving is how humans and organizations, as structured groups of individuals, search for alternatives in the solution space. The academic literature has produced a large amount of work related to this question and numerous researchers have used the NK-model as a way of conceptualizing this search process. The majority of this body of literature, however, examines search through the use of simulations and computational models, which, although oftentimes based on stylized facts of observed behavioral patterns in human agents, do not necessarily represent human search behavior precisely. Only recently, experimental work in this field has gained popularity (Billinger et al., 2014)

Nevertheless, empirical, and especially experimental, evidence on how people search has been scarce. Even more so, when it comes to the question of how groups of individuals, unstructured as well as within a hierarchical order, perform search tasks. Acting collectively has been found to change human behavior in various different settings (e.g. Taylor & Greve, 2006; Mynatt & Sherman, 1975). Furthermore, introducing structure to a group of agents also has been shown to produce numerous different effects, creating new and amending existing biases and tendencies (e.g. Jansen, van den Bosch & Volberda, 2006). Following these findings, it is to be expected that search in (un-)structured groups will differ from search strategies employed by individuals.

In that sense, the present study aims at addressing this issue that, so far, has not received a lot of attention from the scientific community. An experiment is used in order to answer the question if and how people's behavior in search processes changes when they are part of a group, both with and without some form of hierarchy. In the beginning, search as a concept and important aspects therein are presented and discussed in regard to the goals of this work. A special focus is put on the aforementioned NK-model and academic work related to it. Subsequently, the research gap this study aims to address is identified and further crucial aspects, related to groups of individuals and the effects of structure on them, are presented and analyzed. Thereafter, the experimental design is presented, before discussing results and findings of the subsequent analysis. A concluding chapter summarizes and recaps the most important points and discusses a number of limitations of the study.

2 Search in the organizational context

Search tasks appear in numerous situations, including managerial decision making and other organizational contexts, and can be considered a fundamental class of problems present in various fields of everyday life. Search, as the generation of alternatives for solutions to a specific problem, is a highly useful concept in studying human behavior, especially in the realm of organizations and activities related to their functioning.

An organization, at its very core, can be described as a multi-agent system with clear boundaries, which has some form of goal or purpose and where the agents involved contribute consciously to the achievement of that goal (Puranam, Alexy & Reitzig, 2014). The central challenge of organizational design is typically formulated as the need to simultaneously divide labor and integrate efforts in order to achieve the organization's goal (Rivkin & Siggelkow, 2003; Mintzberg, 1979). This produces the need to constantly assess existing solutions and adapt to changing environments and other challenges that may arise and threaten the success of the organization. Reacting to such shocks can be difficult and requires a reliable base of information and knowledge on important aspects that concern the well-being of the firm. In that sense, organizational learning refers to the process of agents acquiring knowledge about relationships between certain actions and associated outcomes (McGrath, 2001). As knowledge is typically limited, at times of only temporary use and the surrounding factors relevant for an organization's functioning can change rather quickly, finding dependable solutions for problems that arise can be highly challenging. Therefore, it can be observed that rather than striving for optimization, organizations engage in search activities, i.e. the generation of alternative solutions to a problem at hand, adjusting decision parameters and eventually choosing the option that yields the best result (Mihm, Loch, Wilkinson & Huberman, 2010).

Search in organizations is ubiquitous and examples for when firms have to engage in search efforts include finding and defining a business strategy (Winter, Cattani & Dorsch, 2007), building organizational capabilities (Gavetti, 2005; Bruderer & Singh, 1996), adapting to industry life cycles (Doty, Glick & Huber, 1993) and developing new products (Mihm, Loch & Huchzermeier, 2003; Yassine, Joglekar, Braha, Eppinger & Whitney 2003; Fleming & Sorenson, 2001). In the realm of innovations, a number of researchers has argued early on that recombination is the ultimate source of novelty (Gilfillan, 1935; Usher, 1954). In that sense, every invention is based either on an entirely new combination of known conceptual or physical materials, or benefits from a newly discovered relationship between previously conjoined components (Schumpeter, 1939; Henderson & Clark, 1990; Nelson & Winter, 1982). This makes the innovative process, a central element to most firms, a fundamentally search-related activity. The extensive interest among academic scholars of management and related domains in this topic of search and concepts such as the exploration-exploitation framework have sparked work on numerous different topics, resulting in a multitude of phenomena being investigated using these concepts (Lavie, Stettner & Tushman, 2010): from product innovation (Tushman, Smith, Wood, Westerman & O'Reilly, 2003), team composition (Perretti & Negro, 2006) and strategic alliances (Lavie & Rosenkopf, 2006) to organizational design (Tushman & O'Reilly, 1996) and organizational search (Billinger et al., 2014). Other managerial tasks appearing in this context have been elaborated on by Siggelkow & Rivkin (2009) who have looked into designing organizational structures, and Laursen & Foss (2003) who examined human resource management practices.

2.1 Search and bounded rationality

In his work on bounded rationality, one of the most important concepts organizational research has built on (Simon, 1955; March & Simon, 1958), Simon (1957) stated that a decision maker typically does not hold complete information on all alternatives for a decision problem but rather discovers those throughout the search process. This has led to the development of two argumentative streams dealing with how organizations make decisions and, more precisely, on which kind of information these decisions are based (Gavetti & Levinthal, 2000). The first line of argument shows a focus on organizational learning, with an emphasis on local search and the development of organizational routines (Levitt & March, 1988). This learning process builds on experiential knowledge and can also be described as a procedure of trial and error learning with a subsequent selection or rejection of certain behavioral patterns (Gavetti & Levinthal, 2000). As this form of information gathering for future decisions is not the only possibility for organizations (March, 1994), the second intellectual stream present in the literature focuses on decision makers explicitly taking into account possible consequences of their actions and anticipating potential implications of certain behavior (Simon, 1991; March & Simon, 1958).

Bounded rationality states that humans have to rely on cognitive representations of problems as they face information-processing constraints which make a certain simplification of the problem space necessary. This becomes manifest in organizational routines and standard operating procedures following experiential learning, as well as in the form of mental models of the environment decision makers develop in order to anticipate consequences of their actions (Gavetti & Levinthal, 2000). Oftentimes, finding new solutions or alternatives resembles a recombination of known decision variables (Becker et al., 2006). Among the most notable in that regard is Schumpeter (1934) who described innovations as the result of recombining resources in a new way (Fleming, 2001; Denrell et al., 2003). Boundedly rational actors, however, cannot process the entire set of available alternatives in a problem space, nor can they fully envision all causal relationships between actions and associated results, as attempts to do so are stymied by the large number of combinations of decision variables and the complexity created by interactions between these variables (Gavetti & Levinthal, 2000).

When it comes to evaluating alternatives, two fundamental modes of assessment can be identified. First, if alternative options can be evaluated based solely on cognitive terms and before engaging in any activity, i.e. the alternative doesn't have to be implemented in order to understand its implications and consequences, the term "off-line" evaluation is used. If, on the other hand, at least a partial implementation of the alternative is required for decision makers to be able to assess its effectiveness, the evaluation mode is referred to as being "on-line", i.e. the knowledge gain is not based on mental models or representations of the problem but rather on the basis of actions actually taken (Lippman & McCall, 1976).

2.2 Search on a fitness landscape

Interpreting the search or problem space of a setting as being a landscape where the searcher has to "wander around" in order to arrive to new solutions is a widely used concept among scholars working on search and related topics. One fairly popular example, which uses the context of firms for illustrating the idea of a fitness landscape was introduced by Levinthal (1997). There, an organization's form is mapped to its relative fitness or likelihood of survival. Building on Wright's (1931) work in the field of evolutionary biology, a fitness landscape is a multidimensional space where each attribute of the firm, or the organism in the original context, is represented by a dimension of the space. Using a final dimension, with the sense of being an "aggregate", the organization's overall fitness to survive is given. If organizations now engage in adaptation and change or modification of a subset of their attributes, they start to "move" over the landscape in search of solutions or configurations of decision variables that yield higher payoffs or indicate a higher likelihood of survival in the future. The height of a given point in the landscape then can be interpreted as the associated fitness level of the corresponding solution (Levinthal, 1997). In that sense, organizations are conceptualized as systems of interdependent choices (Porter, 1996; Whittington, Pettigrew, Peck, Fenton & Conyon, 1999; Milgrom & Roberts, 1995). Hence, in order to understand the performance of an organization, it is crucial to analyze it as a system of interconnected choices (Siggelkow, 2001).

Besides the relatively straightforward issue of facing a large number of possible combinations even when only a few decision variables contributing to the final outcome of a solution are involved, interactions between single choice variables are what drive the complexity of many problems (Levinthal, 1997; Page, 1996). If the contribution of each decision variable to the overall payoff of a solution is independent from all other choices then the complexity of this task is significantly lower (Simon, 1962). When the configuration of one decision impacts the way how other variables influence the final outcome, it becomes much more difficult to identify good combinations. The unpredictability of changes in existing solutions makes experimenting a much riskier endeavor and simply optimizing decision variables sequentially without taking into account that other contributions might change as well is bound to be an underwhelming strategy. Within the subject of business strategy, for example, the challenge can be described as the need to set large numbers of choices in such a way that a competitive advantage can be gained in the environment the firm operates in (Porter, 1991; Rivkin, 2000; Ghemawat & Levinthal, 2008).

When the performance or fitness contributions of individual decisions in an organization are independent of each other, then this firm is said to operate in a simple environment, whereas if decision variables are highly interdependent and exhibit performance-impacting interactions, the environment is considered to be complex (Siggelkow & Rivkin, 2005). In terms of a fitness landscape's topology, this is presumably the most important aspect, its smoothness or ruggedness. It describes the change in the fitness value or performance when one attribute or decision variable is changed. One example could be when a firm changes its strategy then this might lead to significant changes in the performance or fitness level even if all other choice variables are held constant. This indicates discontinuity or ruggedness with respect to one dimension, in this case the choice of strategy. If, however, the value of a particular attribute interacts with other features of the organization and determines or at least influences the fitness contribution of other decision variables, the landscape will show multidimensional ruggedness and will tend to have numerous peaks and few smooth areas. In such a case, even if only one attribute of the firm is changed, without adjusting any other elements, the overall fitness level can change drastically as the singular adjustment has effects on multiple other fitness contributions within the organization. It is usually this interdependence of choice variables that accounts for the complexity and ruggedness of a fitness landscape (Levinthal, 1997).

2.3 Scope of search – Exploration vs Exploitation

For assessing the search behavior of an agent or entity it is crucial to define ways of how to measure this particular kind of behavior and which dimension or aspects to include in the analysis. A central concept in the field of search behavior is the idea of scope of search. It captures the notion of analyzing search efforts based on the principle of "distance" between, for example, a searcher's current position and an alternative location on the landscape. Based on aspects such as the similarity of the two options, statements can be made about how far or close a certain solution is to another one. Extending this idea of search distance to the usage of existing knowledge or solutions to problems, the exploration-exploitation framework was developed (March, 1991). In the context of inventions, local search, which March (1991) terms exploitation, refers to a process of recombining familiar sets of technology and refining previously used combinations of components (Fleming, 2001). The inventor can benefit in such a situation from existing knowledge and thus exploiting known solutions to a problem. In terms of topography of a fitness landscape, search happens in close proximity to familiar locations in the search space. Distant search, on the other hand, which is referred to as exploration in March's (1991) work, occurs as the importance of known combinations in new configurations of decision variables decreases. Inventors try out entirely new and unknown components or recombinations of those and, thus, explore areas in the search space where less knowledge is available about solutions located there, which corresponds to such alternatives being farther away on the landscape than only slightly modified existing options (Fleming, 2001). The neighborhood of a location on the landscape, i.e. a group of points in close proximity to each other, is defined as the solutions or configurations of decision variables that differ from each other in only one attribute. This means that every point in a fitness landscape consisting of N elements, has N points in its immediate neighborhood (Levinthal, 1997).

Where a disproportionate focus on exploration can lead to suffering from the costs of innovation without gaining many of its benefits, and excessive exploitation can trap firms in non-optimal positions, it becomes apparent that maintaining a reasonable balance between those two activities is necessary for the long-term survival and prosperity of an organization (March, 1991). If resources are limited and required in both explorative and exploitative activities within an organization, as they oftentimes will be, then the two concepts can be regarded as, in principle, mutually exclusive, i.e. one can only be increased if the other is decreased, and, thus, as spanning a continuum or spectrum between them (Gupta, Smith & Shalley, 2006). The dilemma arising in allocating resources to either exploitation or exploration activities in an organization manifests itself both explicitly and implicitly: explicitly, in the form of calculated decisions regarding competitive strategies and investment behavior, and implicitly, by choosing organizational structures and customs, ranging from rules and practices for dealing with slack or targets to the type of incentive systems implemented (March, 1991).

A scenario where a firm faces the need for strategic renewal, e.g. due to major changes in the environment, illustrates the constant tension present in the trade-off between exploitation and exploration: on the one hand, there is the necessity to constantly exploit existing competencies in order to maintain stability and continuity, and on the other hand, a simultaneous need for exploration of new alternative solutions arises, crucial to preserving competitiveness and flexibility in cases of radical changes in the business environment (Floyd & Lane, 2000). Both concepts are vital for the well-functioning of an organization and are typically regarded as being inseparable (Levinthal & March, 1993). Exploratory learning is critical to an organization's ability to create variety and adapt to changes that require modification of organizational aspects (McGrath, 2001). In cases when important underlying technologies change, incumbent firms in affected industries who struggle with the arising need to adapt due to an inability to engage in appropriate innovative activities, face serious hardships up to the point where the survival of the organization is at stake (Tushman & Anderson, 1986). In changing environments, it can be crucial for firms to not only focus on things already done well, i.e. put an emphasis on improving mean performance, but to also engage in activities that foster variance in performance, namely, investing time and resources into exploratory innovation as exploration in innovative efforts is typically associated with enhancing variance within the organization (McGrath, 2001). Nevertheless, it can be difficult for organizations to adapt to changing environmental conditions as, even if a firm holds all necessary skills, decision makers within the organization can find themselves restricted in their actions by historical precedents or organizational politics and routines already established (Burgelman, 1994).

Although oftentimes in parts of the literature exploration and exploitation are treated as very different, if not opposing, concepts, both can be regarded as practices of combining existing knowledge, either in new and innovative ways, which would translate to exploration, or in well-established and understood ways, constituting exploitational behavior. In that sense, it is less the difference in the kind of knowledge that is used for either activity but rather the way of how it is used, and the type of goal that should be achieved (Taylor & Greve, 2006). As local search makes use of existing knowledge for purposes which share significant similarities with the current situation, one might expect that it is more certain and more promising than engaging in risky distant search efforts. It might not offer the potential of a radical, highly impactful innovation but will likely limit search related costs and failures due to unsuccessful solutions (Fleming, 2001). Learning from failures is, furthermore, an important aspect in the ability to draw from existing knowledge in local recombination processes. Choices that have led to positive outcomes will be reinforced, whereas actions that have been linked to negative or detrimental results will be avoided (Gavetti & Levinthal, 2000). Searching agents learn which combinations and configurations turn out to be less successful and, thereupon, avoid these, resulting in a process of distinguishing and secluding areas in the search space that are less attractive (Vincenti, 1990).

In Levinthal's (1997) work, computational search agents, which can be thought of as organizations who engage in local search, for example, screen their immediate neighborhood for solutions that are superior in terms of fitness compared to their current position (March & Simon, 1958; Cyert & March, 1963). If such a solution, which implies changing only one decision variable at a time, is found, then the organization adapts accordingly and moves to this new position in the landscape. The height of that point reflects the fitness level the search entity obtains by applying the corresponding configuration of its attributes (Kauffman & Levin, 1987; Smith, 1970). It is important to note that in this model agents do not strictly choose the best performing solution among all neighboring alternatives but rather follow the first one that is discovered to yield a better fitness level than the current position. In contrast to local or narrow search, organizations can also perform what Kauffman (1993) terms as "long jumps" which refer to agents adopting solutions far away from their current position on the fitness landscape (March & Simon, 1958).

In the context of technological innovations, Stuart & Podolny (1996) describe local search by organizations as initiating research and development projects that share certain technological aspects with outcomes of prior projects or searches. This captures the notion of search being "local" if the new solutions or alternatives share a lot with the existing configuration of decision variables. The greater the similarity between two options, the "closer" they are to each other and, thus, the more local the associated search activity is. Specifically, the authors introduce a different measure to assess "localness" of search and distances in search efforts between organizations than what is used in other models. They analyzed patent data on Japanese technology firms and mapped the sector using information about which companies built their innovations and innovative capabilities on which particular existing technologies. By constructing this history- or path-dependent assessment of an organization's current position in the industry, they were able to compare different firms to each other and, furthermore, estimate the distance in this "technological network" between single companies. Firms who based their technological advancements on a similar set existing knowledge were regarded as being close to each other, i.e. similar in terms of structure and innovative role, whereas organizations that had founded their expertise on different patents and technologies were classified as competitors with a larger distance between them. Thereby, the authors were able to specify network distances by varying degrees of structural similarity.

In this context, local search can be interpreted as a company expanding its innovative capabilities into areas where expectedly similar firms, in terms of structural and role equivalence, will be either operating already or, with a certain likelihood, move there in the near future (Stuart & Podolny, 1996).

3 The NK model

A way of modelling interactions in combinatorial search tasks that has become quite popular in the field of management studies is the so-called NK model, initially developed by evolutionary biologist Stuart Kauffman (Billinger et al., 2014; Kauffman & Weinberger, 1989). In that framework, an entity is characterized as consisting of N attributes where each single attribute can take two possible values (Kauffman, 1993). Different configurations of these attributes then represent different variations of the entity, which can be thought of being a combination of genes in an organism, a way to organize a firm or any type of solution to a problem consisting of more than one binary decision variable. The following chapters discuss the main aspects of NK models and focus on highlighting important properties of it.

4.1 Construction of a fitness landscape

The NK algorithm creates a fitness landscape which is a multidimensional mapping of all possible combinations of attributes or parameters. In the context of search, this can be thought of as a task or problem where for every solution there is a corresponding payoff, performance or fitness attached. Depending on the level of complexity, determined by the extent of interaction between single decision variables, the landscape generated can be more or less rugged. As individual choice variables are always binary in the NK model, there are 2^N possible configurations of the entity, which constitutes the entire search space (Levinthal, 1997). The total number of choice variables is denoted by N. The parameter K reflects the degree of interdependence between decision variables and determines the ruggedness of the landscape. More precisely, K is the number of variables that contribute or influence the payoff of one single choice variable. If K = 0, then every variable's contribution to the overall payoff of a solution is unaffected by the values of all other decisions. This means that for every variable there is a clearly superior configuration that will maximize total payoff and whichever of the two possible states a single decision variable can have, increases final profit by a larger amount should be implemented. On the other end of the spectrum for values of K is K = N – 1 which would indicate that the contribution of a single choice depends on every other decision variable present in the task (Billinger et al., 2014).

An illustrative example on how the NK model works is given by Levinthal (1997). Consider an organization or, more generally, an entity with 10 different attributes, so N = 10, and a complexity level of K = 3. Suppose that this entity is currently specified by the following string: 1, 0, 0, 1, 1, 1, 0, 1, 0, 0. Further, assume that the three attributes influencing the fitness contribution of every choice variable, given by the value of the K parameter, are always the following three elements in the string after the particular decision variable. So, in this example, the fitness contribution of the first parameter depends not only on its own value, which is currently 1, but also on the values of the second, third and fourth element in the string, currently showing values of 0, 0 and 1, respectively. The string is assumed to be circular so for elements close to the end of the string the relevant influential attributes are those at the beginning, e.g. the fitness contribution of the ninth element is dependent on values of the tenth, first and second feature. For every possible configuration of values for each attribute a random number ranging from 0 to 1 is drawn from a uniform distribution and constitutes the fitness contribution of this attribute for the given constellation of own value and values of all other relevant choice variables. To give an example, that means that for the case of the first element holding a value of 1 and the second, third and fourth element, which are the relevant ones for determining the fitness contribution of the first attribute, showing values of 0, 0 and 1 respectively, a random number between 0 and 1 is drawn. For the case that the value of the first variable is changed to 0 and all other values stay the same, a distinct random number is drawn. Then the value held by the variable in the second place of the string is changed, another different random number is drawn and so on. This process is repeated for all of the 2^{K+1} different combinations for each attribute and for all N attributes in the string. The overall fitness or performance level for a particular configuration of all N features is then calculated as the average of all single contributions by individual decision variables (Levinthal, 1997).

4.2 Structural properties of the model

If K is zero, then the corresponding fitness landscape shows only peak with all surrounding solutions gradually and smoothly ascending towards it. In this case, the maximal difference between two adjacent points on the landscape, i.e. two configurations or solutions which differ in only one attribute, is 1/N, which is the highest possible contribution of a single choice variable. As K increases, however, and the fitness landscape becomes more rugged, changes in single features affect other elements' fitness contributions as well, resulting in the maximal difference between two neighboring points now being (K + 1) / N. In the extreme case of K = N – 1, a single adjustment in an existing configuration, which results in changing fitness contributions for all other decision variables as well, can lead to drastically different levels of fitness between solutions in close proximity of each other, making the landscape extremely rugged (Levinthal, 1997). For K = 0, however, the fitness contribution of every individual choice variable does not depend on the configuration of the rest of the attributes, which implies that every element can be set to the value that yields maximum payoff without impeding other contributions. In that way, it is possible to gradually ascend to the global peak and reach maximum fitness levels in a relatively easy manner. This process is usually referred to as "hill-climbing". Neighboring solutions of the status quo are examined and if any of these alternatives shows to bear a higher payoff than what can be achieved with the existing configuration, the searcher moves to this new alternative. The process is then repeated until no neighboring alternatives can be found that yield a better payoff. In this case, a peak is reached, and narrow search can no longer achieve better results. The peak, however, can be a global one or merely a local maximum (Billinger et al., 2014).

In a scenario where K is very high, narrow search is not likely to lead the searcher to the global maximum. The landscape is rugged, meaning that there are numerous local peaks where simple hill-climbing will terminate the search process prematurely. As changing one decision variable now not only directly changes total payoff but also affects the contributions of all other choice variables, it becomes very difficult to find high-performing solutions by only exploring alternatives in close proximity to the status quo (Billinger et al., 2014). In this setting, a singular change in decision variables can lead to a decline in overall fitness levels although the same change in combination with adjustments on other features can turn out to be beneficial in terms of fitness or performance. This shows that through changing potential differences in fitness levels between neighboring points in the landscape and an increasing number of peaks throughout the search space, the K parameter determines ruggedness and, consequently, complexity of a search environment (Levinthal, 1997). As stated, with increasing levels of K, the number of local peaks in the landscape increases as well (Rivkin, 2000). Where with K = 0 there is only one single peak in the landscape, K = N – 1 usually results in around 2N / (N + 1) of the 2N possible alternatives being local peaks (Kauffman, 1993). Furthermore, with increasing K, the average height of local peaks tends to shrink, up to the point where with K = N – 1 the value of a typical local peak drops to 0.5 which equals the expected value of a randomly selected point in the landscape (Rivkin, 2000). At lower levels of K, the situation for search agents is not quite as underwhelming as in the extreme case but still a local peak can be expected to show a value of around 0.7, which, at least for large values of K, corresponds to mediocre performance (Weinberger, 1991; Rivkin, 2000).

Every peak in a fitness landscape has what is called a basin of attraction. A basin of attraction refers to the area around a peak for which local search efforts are likely to result in the search entity reaching the corresponding local maximum and, if no distant search or long jumps are possible, the search process will be terminated there (Kauffman, 1993). A notable property of these areas is that they tend to be positively correlated to the height of the peak they are surrounding (Levinthal, 1997). In other words, higher local maxima tend to have larger basins of attraction. As the complexity in a landscape increases and it becomes more rugged, average peak height tends to decrease. This, in turn, implies that the basins of attraction related to those maxima also tend to diminish which makes it far more difficult to perform successful long jumps and land in the basin of attraction around a promising local optimum. Considering the global peak in a landscape, which is usually surrounded by one of the broadest basins of attraction, an example illustrates just how quickly a rise in the number of decision variables can render global optimization more challenging if the landscape is very rugged: for a scenario with 8 choice variables and maximum variable interaction, i.e. N = 8 and K = 7, the global peak's basin of attraction covers around 7% of the entire search space. After an increase in decisions involved and complexity remaining maximally high, i.e. N = 16 and K = 15, the basin corresponding to the global optimum reduces to a mere 0.5% of the landscape (Rivkin, 2000).

Another notable effect that a rise in the complexity parameter K has on the topography of the fitness landscape is that peaks tend to spread all over the search space, i.e. the distance between local maxima increases with the interdependency between decisions. Rivkin (2000) shows that for low levels of K, there is a negative correlation between a peak's height and its distance to the global optimum. In other words, the higher the local peak, the closer it tends to be to the global one. With growing levels of complexity, this relationship fades away, and peaks spread apart, resulting in the location of a high peak carrying less and less information about where other attractive solutions can be found. The intuition behind this observation builds on the notion that as decisions show low levels of interdependency between each other, the strategy of a well-performing company, for example, can usually be decomposed into groups or subsystems of choices that work well in a certain configuration without reacting too much to the arrangement of the remaining choice variables (Simon, 1962). Consequently, adjustments within such subsystems typically don't affect other parts of the strategy which, in terms of landscape topography, can be interpreted as high peaks being located close to one another as high performing solutions only differ from each other by relatively small modifications within subsystems of choices. In a practical context, this means that firms which exhibit high levels of performance, are likely to show similar configurations in the majority of decision subsystems, i.e. they will all follow a certain "theme" with slight modifications within groups of choice variables. If, however, interactions between individual decisions are more pronounced, i.e. the parameter K shows higher values, an organization's strategy cannot be separated into relatively independent subgroups of choices which, in turn, suggests that firms on higher peaks probably operate on different ways of business altogether, meaning they are expected to be found farther away from each other than in low complexity scenarios. The global maximum, in this case, is less likely to provide a model or template for a successful business strategy suitable for adaptation, as it carries only little information on other well-performing configurations. Hence, equally effective organizations may differ significantly from each other in terms of how they achieve high performance (Rivkin, 2000).

4.3 Theory of NP-completeness and intractability of the model

When facing a decision or search problem represented by an NK model, one might be tempted to think about ways to "solve" problems like this mathematically or computationally, in the sense of screening all possible combinations of choice variables and their associated performance levels quickly and then simply selecting the best-performing one. Alternatively, and more realistically, one could think of using existing solutions and available knowledge to swiftly identify attractive peak locations. That this can turn out rather quickly to be a practically impossible task, is shown in Rivkin's (2000) work on imitating successful strategies.

The author first proceeds to show that global optimization becomes very difficult even with relatively low levels of K. For this, he uses the theory of NP-Completeness (Johnson & Garey, 1979). The theory of NP-Completeness basically deals with sets or classes of problems and how easily they can be solved. In other words, the solvability or "tractability", as it is called in the respective literature, is given by a function which relates the size of a given problem to the time it takes to solve this problem using an algorithm, with an algorithm being a general, step-by-step procedure for solving problems. Two types of algorithms can be identified along the lines of this argument: polynomial time algorithms and exponential time algorithms. A polynomial time algorithm is one where the time function is some polynomial function which means that as the size of the problem the algorithm is solving increases, the time required for getting to a solution increases linearly. If the time function is an exponential function, which is the case for exponential time algorithms, then the time associated with solving a problem using this algorithm increases a lot faster, i.e. exponentially, with growing problem size. The importance of this distinction lies in the aspect of "increasing a lot faster". Exponential time algorithms become intractable, or unfeasible, very quickly and with relatively small problem sizes compared to polynomial time algorithms.

This means that for large sizes of problems, algorithms with exponential time functions effectively become "unusable" in the sense that it would take an impractically large amount of time to solve the problem with that algorithm. If, then, a problem can be solved using a polynomial time algorithm, it is considered to be part of the P class of problems. For problems where no such algorithms exist and only exponential time algorithms can be used, the NP class has been defined, which relates to the notion that in order to solve such a problem in polynomial time, one would need a nondeterministic machine, i.e. a computer with unlimited processing power. Furthermore, the NP-complete class of problems, a subset of the NP class, contains problems that are at least as difficult to solve as any other NP class problem and, additionally, can be transformed into other NP-complete problems in polynomial time. This means that every NP-complete problem can be converted into any other NP-complete problem. This, in turn, implies that if a polynomial time solution to any NP-complete problem could be found, the entire class of NP-complete problems could be solved by simply transforming every problem into the specific solved one. However, to date, no polynomial time algorithm has been discovered to solve any of the NP-complete problems and many experts in the field suggest that all NP-complete problems are indeed intractable, i.e. (for large sizes) not solvable with any kind of algorithm in a reasonable amount of time.

Linking these insights to the NK framework, Rivkin (2000) shows that for fitness landscapes under the NK model two scenarios arise: as long as the complexity parameter K is sufficiently low, global optimization is algorithmically solvable as a P class problem. In the extreme case of $K = 0$ this would simply mean changing decision variables one by one until the global maximum in the landscape is reached. However, as soon as K exceeds two, i.e. for all landscapes with $K > 2$, finding the global peak turns into an NP-complete problem and the associated optimization problem has thus to be considered intractable. For a search task to be truly "practically unfeasible" though, the author suggests two additional conditions must hold: the number of decision variables present must be substantial and the time it takes to evaluate one alternative in the search space must be non-trivial. If, in fact, the parameter N and evaluation time per solution are low enough, a simple exhaustive search algorithm, screening all possible alternatives on the landscape, could be employed.

Although this kind of search procedure does not depend or react to complexity, i.e. the parameter K, a simple example illustrates just how quickly a search task exceeds reasonable amounts of processing time: consider the case when $N = 5$ and the evaluation time per alternative, which can be thought of as the time it takes to estimate the performance or fitness level of an individual configuration, is one hour. With $2^5 = 32$ different possible combinations of choice variables, it would take 32 hours to evaluate every point in the search space and find the global maximum. If N is increased to 10, however, total processing time increases to 43 days. Raising the number of decisions to be taken to 20, results in the exhaustive search algorithm needing 120 years to cover the entire landscape. The reason for this rapid increase naturally lies in the model's property that the number of possible combinations grows exponentially with an increasing total of decision variables. In summary, global optimization on NK landscapes is particularly difficult mainly due to two aspects: first, the fact that search space grows exponentially with increases in the number of decisions involved, which renders exhaustive search strategies unfeasible fairly quickly, and second, the aspect that even with faster and more sophisticated algorithms, finding the global maximum in a landscape with high enough interaction between decision variables, i.e. if $K > 2$, turns out to be a NP-complete problem and is thus not solvable by any kind of algorithm in a reasonable amount of time.

4 Literature review

This section aims at analyzing existing academic literature on search and identifying prevalent propositions on forms of search behavior related to individuals, groups and structured groups, i.e. organizations. A special focus is put on empirical and, especially, experimental evidence available on search processes. The section concludes with identifying the research gap arising and how this study aims to address the related issues discovered.

4.4 Individual search

One important point has to be made right at the beginning of this chapter, which is that although many of the studies presented in this work and, especially, in this chapter talk about organizations, firms and companies, most of the concepts analyzed are based or relate to individual agents. In that sense, when the behavior of an "organization" is described, e.g. in one of the numerous simulation studies, the phenomena translate to individuals, as the decision makers are single entities characterized by only one "actor" making the decision. This actor can then be regarded as a firm, in the sense that it acts and behaves as a single agent, or as an individual in the actual sense of the word. Studies where more than one entity is involved in the decision, will be discussed in the subsequent chapter.

Although, authors dealing with technological change in the management literature, and among those prominently proponents of the bounded rationality approach (Fleming, 2001), have characterized the underlying search process of such change as uncertain and more or less blind, as in not having any information available on the properties of the corresponding search space (Vincenti, 1990; Nelson & Winter, 1982; March, 1991), most firms in the real world will likely have some relevant information on their environment and will thus be able to perform some kind of "informed" search rather than simply navigate the search space blindly looking for high-performing solutions. One, rather straightforward, source of information that can be used in the decision making or search process is existing knowledge or available information that the agent holds. In a very fundamental form, this knowledge can be understood as the searcher's current position in the search space and the corresponding fitness level. Based on this information, search entities are already equipped for performing local search, as this option relies most heavily on the current situation or location on the landscape.

According to Stuart & Podolny (1996) the local search behavior that can be found to be predominantly the case in technology-oriented firms is history or path dependent and results from processes at the individual and organizational level in a firm, as well as from the very nature of how the firm's innovative capabilities were built. For boundedly rational, individual decision makers the sheer number of possible alternatives in terms of where to direct R&D funding is a major issue contributing to the observed tendency towards local search. In a situation with numerous decision alternatives, it appears reasonable to turn to previous decisions for guidance (Stuart & Podolny, 1996). Additionally, as the environment the organization operates in is typically characterized by uncertainty regarding technical, social and economic factors, and, therefore, also the anticipation of the competition's behavior is usually relatively difficult, decision makers are likely to rely on past experiences and historical precedents when it comes to setting the path for future innovative activities (Tushman & Rosenkopf, 1992; MacKenzie, 1992). Over time, if the nature of the firm's habitat persists to display high levels of uncertainty for critical factors, organizational norms form which tend to reinforce this behavioral pattern and hinder the emergence of alternative, potentially more risk-seeking and exploratory habits among managers and decision makers (March, 1988). In that sense, results from prior search efforts, i.e. established technological or innovative capabilities, act as natural starting points for further innovation and subsequent search activities (Nelson & Winter, 1982). The process of ongoing local adaptation also provides an important implication for understanding the existence of diversity in organizational forms. Namely, the fact that an organization's possibilities in adapting and modifying its structure will heavily depend on the initial "starting position" of a firm on the fitness landscape. Based on that, which in reality can be interpreted as the history of the firm or the type of structure that was chosen at the moment of founding, local search will be limited to certain areas in the whole search space, meaning that there is a limited number of configurations that a firm can reach within a certain amount of time. This approach can be regarded as an extension to existing explanations to why firms differ in terms of their organizational structure, such as contingency theories or ecological arguments (Levinthal, 1997).

Another source of information might be the competition active in the same industry or area. In that sense, organizations might decide to engage in imitation of their competitors' ways of doing business in order to profit from already existing knowledge and experiences and thereby avoid costs involved in search processes. When it comes to barriers to imitation, i.e. reasons why imitators don't or only partially succeed in their efforts, large parts of the existing related research can be classified into two groups (Rivkin, 2000): the resource-based view of the firm focuses on aspects such as tacit knowledge, causal ambiguity, social complexity or first-mover advantages that make it difficult for competitors to copy a successful firm's strategy. Industrial economists on the other hand put their emphasis on actions taken by incumbent firms in the industry specifically aimed at deterring copycats from attempting imitation, which can range from retaliation threats to various forms of costly commitments that may even influence their own future incentives. Rivkin (2000) himself offers an additional explanation to why some companies manage to stay unmatched for long periods of time even if the main ingredients of their strategy are publicly known. The argument is based on the very nature of search tasks and especially the complexity therein. Recall, that the complexity of a business strategy is driven by the number of decisions that constitute the strategy and the degree of interaction that occurs between individual decisions (Simon, 1962). In that sense, the author argues that oftentimes the mere complexity of a business strategy might be detrimental enough to competitors' efforts to imitate it. When a large number of decisions or relevant aspects is involved and factors constituting the strategy show high degrees of interaction and interdependency, attempting to copy the solution in its entirety is by all means a daunting task. Additionally, even if imitators attempt to reproduce only parts of the overall configuration, the chances of success quickly dwindle with increasing complexity of decision interdependencies, as on very rugged fitness landscapes adapting only subsets of configurations or even getting a small portion of them wrong bears the potential for ending up with solutions that perform far worse than possibly even the initial starting point before the imitation attempt (Rivkin, 2000).

Levinthal (1997) implements long jumps as search agents drawing a random configuration of decision variables in every round, comparing the associated fitness level with their current position and, if the newly found alternative is superior to the status quo, adopting the new form. In other words, the organization, consisting of N attributes, changes its structure randomly to one of the 2^N possible alternatives present in the entire fitness landscape, given that relocation to the new position yields better results. A few interesting observations in regard to this process of innovative search are also noted by the author. First, as the likelihood of a firm finding an attractive solution through innovative search depends on its current fitness level, the relationship between poor performance and increased probability of performing long jumps emerges endogenously in the model. If an organization is located on a point with high relative fitness, then it is unlikely for this firm to randomly find a position on the landscape that holds a better solution than its status quo. Therefore, even if successful organizations engage in the same amount of innovative search efforts as those who are worse off in terms of fitness levels, high-performing search entities are far less likely to actually employ radical changes compared to agents currently facing less favorable outcomes. The second observation deals with the relationship between the rate of long jumps and local adaptation over time. For this, it is assumed that search entities can search both for distant alternatives as well as adapt locally. In either case, the rule of only moving to a point that is superior to the current position remains unchanged. At the beginning of the search process, most agents will likely show low to medium fitness levels as they are released randomly throughout the landscape. Accordingly, the probability of finding superior solutions by chance, i.e. through long jumps, is higher at this stage than later on. As firms proceed to occupy alternatives that are situated higher on the fitness landscape, it becomes more and more unlikely to find better performing configurations through innovative search. Thus, as time passes, long jumps become increasingly rare and local adaption can be observed more often to yield favorable results. Additionally, if a relocation to a distant position turns out to be beneficial, it is likely that this change bears the potential for further local adaptation. The firm might move to an entirely undiscovered area in the search space where there is room for improvements through local search efforts. Hence, as successful long jumps can potentially renew opportunities for local adaptation, the combination of distant and local search appears to be a promising strategy (Levinthal, 1997).

Siggelkow's (2001) work on how firms adapt in changing environments provides a number of interesting observations on the situation an organization faces when it is stuck on a local peak. Basically, the company has three options: a) stay where it is, b) move away with small, incremental changes, i.e. engage in local search, or c) try to rearrange numerous organizational aspects at once to find a new peak, i.e. opt for distant search and attempt a long jump. The first option clearly doesn't change anything for the firm as long as the underlying landscape, i.e. the environment the company is operating in, doesn't shift. Engaging in local search may appear to be a reasonable option as the adjustments remain manageable and recombinations of decisions mainly build on the existing solution and available knowledge. Nonetheless, decision makers within the organization might argue that although search related costs and the risk involved in this approach are limited, incremental change will very likely lead to a decrease in performance. The intuition for this is straightforward: the existing solution, i.e. the current position on a local peak in the landscape, builds on a fully aligned system of relevant decisions which was already optimized, in all likelihood by incremental adaptation, to achieve the highest possible performance or fitness level in that particular area of the search space. The firm is stuck on the local peak and held there by its existing configuration, in other words, it has fallen into a competency trap (Levitt & March, 1988; Levinthal, 1992). This leads to the conclusion that the only possibility for the organization to overcome this sticking point and leave the local peak, is to perform a long jump and try to locate a higher peak farther away on the fitness landscape. Such an endeavor, naturally, is much more challenging and typically bears larger associated risks as it requires a large and complicated rearrangement of multiple aspects of the organization and only pays off if the firm manages to end up on a higher peak or, more generally, point with higher fitness values on the landscape (Siggelkow, 2001). An aggravating factor can be the circumstance that a specific target, usually a peak, may be unavailable or unknown. And even if there is a clear goal of where the organization wants to relocate itself to with a long jump, for highly complex and interactive landscapes with formidable size this might still prove to be a highly demanding task. With a lot of decision variables, namely, which translates to a high value of N, and high interdependency between those decisions, i.e. significant levels of K, a distant search approach that is even a little bit off, can end up leaving the firm even worse off than before. More precisely, as N increases, the likelihood of getting some of the relevant decisions wrong rises, whereas with growing K, errors tend to matter more, as one suboptimal decision can drastically hamper the beneficial effects of other aspects in the system (Rivkin, 2000).

An illustrative example from a practical point of view which shows just how difficult it can be to opt for incremental adaptation when trying to leave a local peak, that possibly has turned into a sticking point, comes from the U.S. automobile industry. American car manufacturers observed that Japanese competitors achieved remarkable performance levels due to their lean manufacturing techniques and decided to adopt some of the aspects of this strategy. However, by only adjusting relevant factors punctually, such as investing in flexible machinery (Jaikumar, 1986), without taking into account the high degree of interdependencies present in this manufacturing system, these single-dimensional changes were not able to contribute greatly to improving performance. Only after approaching the restructuring of the production process in its entirety, dealing with issues such as workforce training, inventory practices and supplier relationships, companies attempting this adaptation challenge were able to successfully implement the lean manufacturing strategy and reap the benefits of the new techniques (Siggelkow & Rivkin, 2006; Milgrom & Roberts, 1990; MacDuffie, 1995).

The decision on where on the exploration-exploitation spectrum to set a firm's innovative activities, largely depends on the type of environment the organization is operating in. In dynamic settings, exploratory innovation appears to be more effective, whereas in strongly competitive settings, a more exploitative approach to innovative action is appropriate (Jansen et al., 2006). In this trade-off between exploration and exploitation or on the continuum that those two concepts generate, many scholars agree that striking the right balance between the two approaches is crucial. As balancing the two implies that both have to be pursued more or less simultaneously, an organization faces requirements to enable and facilitate operations for both means. One option to do this is via organizational separation which basically captures the idea of assigning different parts or divisions of the organization to either explorative or exploitative activities. Thereby, the firm as a whole can benefit from both approaches, whereas the responsible subunits can focus exclusively on their assigned tasks. Other possibilities to structure frictionless work for balancing exploration and exploitation are temporal separation and domain separation (Lavie, Stettner & Tushman, 2010). As the complexity of a search space, in other words, the topology of a performance landscape, depends on the degree of interaction and interdependence of choice variables, in the context of organizational structure, this translates to various decisions such as product lines, production strategies and other aspects showing varying levels of interdependence across firms (Levinthal, 1997). Among the more prominent pieces of research in management literature on this topic is Chandler's (1962) work dealing with the relationship of a firm's strategy and its organizational structure. This notion was further developed into ideas of the overall configuration of an organization's strategic choices determining or at least contributing to the development of the firm's structure (Miller & Friesen, 1984).

The exploitation and exploration framework can also be used to explain, at least partly, why interdependencies between elements of organizational design can be observed. Design elements, namely, influence how broadly an organization searches when looking for attractive solutions in the decision space and, additionally, affect the firm's ability to stabilize around such a point in the search landscape as soon as one is found. Simulations show that interactions between various elements in an organization, such as the type of CEO, restrictions on the flow of information, incentive schemes and task decomposability, result in different firm-level behaviors regarding exploratory efforts (Rivkin & Siggelkow, 2003). In a highly complex environment, represented by a very rugged fitness landscape, a group of dominant organizational forms is likely to emerge from the total population of firms operating in that particular market or area. Searching entities travel over the landscape in search of peaks, corresponding to dominant configurations of choice variables in regard to the local neighborhood. This ongoing process of narrow improvements and adaptation results in a systematic heterogeneity of organizational forms and the emergence of few firms occupying local maxima in the confined action space, niche or market. In that sense, the presence of complexity due to decision variable interdependency provides an explanation for existing variation between organizations, even if they operate in the same environment, and the occurrence of dominant structures or forms among the overall population in that environment (Levinthal, 1997). Two distinct arguments can be found in the literature to motivate the occurrence of radical changes in organizational structures (Levinthal, 1997). One relates to the idea that firms engage in activities that foster innovation and that these efforts can, with a certain probability, pay off and lead the organization to new ways of doing business or new forms of technology that consequently increase the firm's fitness level and performance (Nelson & Winter, 1982). The other line of argument revolves around the notion of failure-induced search which describes the tendency to shift from incremental improvements to more radical changes in situations where performance is declining or the overall well-being of the organization is in danger (Tushman & Romanelli, 1985).

4.5 Group and Structured-group search

Search processes are oftentimes performed by collectives or groups of multiple agents simultaneously participating in the generation and evaluation of alternatives. In the context of organizations, which can be regarded as structured groups themselves when observed in their entirety, numerous decisions are made not by just one single member of the organization, but regularly dependent on multiple actors which interact in teams, share decision rights or exhibit some kind of hierarchical relationship between each other. Structure and hierarchy are especially interesting for management scholars and have been the focus of a multitude of academic pieces of literature. These works and contributions from various different fields of research have revealed that decisions made in groups frequently differ systematically from those resulting from individuals acting on their own. Hence, it is to be expected that search-related behavior will also reveal variation between single agents and groups. A collection of relevant findings is presented and discussed in this chapter.

Group decision making refers to a situation where multiple agents come together to jointly solve a problem at hand (Kabak & Ervural, 2017). This collective of individuals can either be unstructured, i.e. the members of the group are all equally responsible for the outcomes they produce but no special roles or rights are assigned to any individual, or structured, in a sense that there is a certain formal, hierarchical order in place, defining aspects such as decision rights, authoritative relationships or responsibilities for specific portions of the task to be worked on. It has been argued that in collective search and innovation processes the extent to which exploitative and exploratory behavior can be observed depends on both formal hierarchical structures and informal social relations (Cardinal, 2001). Compared to an individual decision maker, a number of agents working together as a group is typically regarded as being able to deal with more complex issues and to arrive to solutions more effectively. The intuition behind this is that a collective, due to differing backgrounds and personalities of members, usually has access to a larger pool of knowledge, can therefore share and process more information and, eventually, is able to evaluate a larger number of alternatives than a single decision maker (Bernstein, Shore & Lazer, 2018).

In terms of structure and hierarchy, research has also revealed that centralization in organizations tends to negatively affect exploratory innovation, whereas increasing formalization typically fosters exploitative innovative processes (Jansen et al., 2006). For firms who find themselves in a situation where the necessity to adapt arises, e.g. due to a shift in the environment, Siggelkow & Levinthal (2003) find support through computational models for the use of temporary decentralization. Although a trade-off arises between short-term costs of decentralized exploration and the associated long-term benefits of finding higher performing configurations of decision variables, the authors show that temporarily decentralized organizational structures outperform purely decentralized and centralized alternatives in the long-run. As the underlying fitness landscape, i.e. the search space, becomes more complex and interactions between decisions and units of the organization increase, the optimal length of decentralized exploration tends to grow (Siggelkow & Levinthal, 2003).

In the context of organizational level behavior, local search can be regarded as the result of smoothly functioning organizational routines (Stuart & Podolny, 1996; Cyert & March, 1963). Following Nelson & Winter's (1982) argument that routines can be described as patterns of activities that are repeatedly used as responses to similar situations the firm faces, the establishment and reinforcement of such behavioral patterns fosters a tendency and orientation towards known solutions and anchors decisions made at the organizational level to existing knowledge and prior experiences, regardless of the particular individual responsible for the decision. Hence, even if individual decision makers were able to overcome their personal bias, as mentioned above, well-established organizational norms and routines will likely restrict them from deviating too much from, usually less risky but also potentially less promising, local search activities (Stuart & Podolny, 1996). In order to build innovative capabilities and the capacity to achieve technological advancements, the organization has to ensure continuous investments in personnel, physical and intellectual property and organizational knowledge relevant to the particular subfield or area it plans to build expertise in (Teece, 1988). If then a certain level of competence is ensured, the firm has a higher likelihood of successfully developing new technologies in that particular area as it can build on existing knowledge and prior experience (Cohen

& Levinthal, 1989). Thus, the resulting concentration of successive R&D activities in areas of established competencies again constitutes a form of predominantly local search, as search efforts revolve around existing solutions, prior experiences and available knowledge (Stuart & Podolny, 1996).

Although collective search has never been observed empirically in the context of the NK framework, a study by Taylor & Greve (2006) shows that team innovations, compared to the results of innovative efforts by individuals, tend to exhibit greater variance in product performance. This observation can be interpreted as an indication of groups of searchers showing a certain propensity for more distant search, as opposed to individual search agents who tend to focus on local search undertakings, even if a broader scope offers greater potential for performance improvements (Siggelkow & Levinthal, 2003). Taylor & Greve (2006) examined data from the comic book industry and found that teams tend to produce higher variances of performance which indicates that they engage in broader and more distant search than individuals. An important caveat to this observation is that the differences in performance were mainly driven by experiences team members hold, such as prior experience of working together or a corresponding record of preceding work in the relevant field. The mere fact of having a larger number of members in the group produced mixed results. Collectives exhibited superior performance due to the ability to draw from the experience of multiple individuals (Taylor & Greve, 2006).

4.6 Experimental studies on search

A large part of academic work on search problems and search-related topics has been using simulation studies to investigate and analyze the effects of factors such as complexity or search agent characteristics on search behavior. Numerous behavioral predictions were produced by the computational models employed but only few have been empirically confirmed or even tested. Among the group of empirical studies in that area of research, experiments have been of rather negligible importance to scholars, which can be seen from the limited number of experimental studies published in the last decades (Baumann et al., 2019).

Drawing on existing work in other research disciplines such as economics and psychology, it becomes apparent that empirical evidence on how individuals approach combinatorial tasks is rare.

While literature from the field of behavioral economics has mainly focused on problems of sequential sampling, where alternatives are given exogenously one after the other, large parts of the existing management literature have worked with the assumption that agents play a crucial role in the generation of alternative solutions. Although psychologists have produced a significant body of work on human behavior in complex and dynamic environments, a frequently missing distinction between task complexity and environmental dynamism can be identified in this particular stream of literature. However, the evidence found in experimental studies relevant to the topic at hand show a tendency towards narrow, neighborhood search and a certain aptness to get stuck, early on, on local peaks. Still, there are important caveats to these findings: the tasks studied in these works were dealing with allocating resources on two alternatives rather than presenting individuals to multiple decision variables where combinations of choices had to be explored in order to find high payoffs. Additionally, interactions between alternatives were either present or not, meaning that there was no possibility of introducing variation in the degree of interaction. Both aspects constitute substantial differences between these studies and the NK framework (Billinger et al., 2014).

Among the most important pieces of literature for this study is the work by Billinger et al. (2014) which introduced the “alien game” as a useful method of presenting human agents a search task with tunably rugged landscapes, following the logic of the NK model. Specifically, subjects in the experiment are asked to use a grid-based interface on a computer to create art for an extraterrestrial lifeform by combining ten different graphical symbols. The ten symbols correspond to the N choice variables, in this case ten, and span up the fitness landscape. The complexity parameter K is varied and acts as the independent variable in the experimental design. Participants have 25 rounds in which they can create configurations to their liking and are presented an associated payoff after every round’s submission. The reason for using the setup of art creation for aliens follows the rationale that searchers should not be able to draw on any knowledge that could influence their search behavior, which is ensured by introducing an entity whose preferences are entirely unknown to players. Thereby, they have to rely on information generated throughout the process of the game and are forced to employ a trial-and-error strategy to learn about the topography of the performance landscape. The main finding of the study is the observation that human agents engage in a form of adaptive search behavior. Depending on the current performance, assessed in relation to a reference point such as the configuration associated with the highest payoff so far, players tend to increase explorative efforts in cases of failures and underperformance, while narrowing search down when solutions are linked to success. The complexity of the fitness landscape had no direct impact on the search behavior but rather influenced the feedback individuals received on their search behavior. In that sense, complexity exhibited an indirect effect on the way subjects played the alien game. The authors additionally provide a computational model in which they implement the empirically discovered search strategy and reveal that adaptive search bears the potential downfall of breaking up local search too early. Economic agents as well as human searchers following an adaptive search strategy tend to not sufficiently screen the local environment of a solution, thus frequently missing out on promising configurations by prematurely switching from local search to more distant attempts. Furthermore, the study was able to empirically replicate and confirm some findings from theoretical work in the NK literature, such as a performance increase in the number of rounds played or decreasing marginal gains of search trials (Billinger et al., 2014).

Mason & Watts (2012) conducted an experiment where they let groups of people perform a search task either as a collective, i.e. part of a network structure where information flow between agents was possible, or as independent searchers who had no such possibility of accessing knowledge based on other individual's experiences. The experiment was conducted online, and the task was to find points on a two-dimensional landscape that would yield high payoffs. By clicking on any point on the landscape participants were presented with the associated payoff of this location. Note, however, that the payoff distribution used was not generated by an NK algorithm, but by an alternative method. Additionally, the search task cannot be described in the logic of the NK model with a number of binary decision variables to be combined. Nevertheless, the setup is worth considering, as well as the associated results. The intuition behind the task is that players are drilling for oil and can only learn how profitable a certain spot is by spending a "try" on drilling at that location. The authors not only differentiate between networked searchers and independent searchers but also define multiple network structures that groups are assigned to. Their findings reveal that, not surprisingly, networked collectives of searchers perform better than independent players as good solutions discovered by any member of a network can spread to other searchers and, thus, increase the overall benefit for the whole group. This, however, creates a social dilemma for individual members of the network as high levels of exploitation of existing solutions promise individual success, because costly exploration is avoided, but simultaneously are detrimental to the collective's overall performance as the likelihood of discovering a high-performing alternative falls with every individual searcher deciding not to engage in explorative efforts. Additionally to this result that largely corresponds to prior findings in the literature, another important observation of the experiment turned out to be quite contradictory to earlier results: instead of inefficient networks, i.e. networks where information, e.g. about alternative solutions, is disseminated less effectively, outperforming efficient structures in situations where substantial levels of exploration are required, the outcomes of the experiment show the exact opposite happening. Prior research, albeit extensively based on simulation studies using computational agents, suggested that inefficient network structures are superior due to the fact that members don't get collectively stuck at local maxima before finding the global optimum or, at least, a high-performing solution (Lazer & Friedman, 2007). As information flows through the network much slower, compared to efficient networks, search agents have enough time to explore the search space on their own and not be tempted to copy any alternatives that might turn out to be low-payoff local peaks. The reason why this expectation was not met by the participants in the experiment turned out to be the fact that searchers in effective networks engaged more in exploration than those in inefficient structures which runs contrary to what computational models would have predicted (Mason & Watts, 2012).

4.7 Research gap

A large portion of the work done with NK models has focused on the use of computational models where the performance landscape created by the NK algorithm is used as the task environment for artificial economic agents. The question of interest in these scenarios is how those agents search the landscape for solutions with high payoffs in a situation where the number of possible solutions vastly outnumbers the available search trials. The specific search behavior, namely the way how agents navigate through the landscape, is usually aimed at behavioral plausibility rather than optimality of search. The stylized behavioral processes these strategies are based on are typically drawn from empirical work on organizational search and adaptation with the behavioral theory of the firm being the predominant source for those (Billinger et al., 2014; Levinthal & March, 1981).

Although the search behavior and the decision rules employed by computational agents are based on empirical insights, there is little to no research on how human agents, as actual decision makers, behave in situations where they face complex combinatorial search tasks and corresponding performance landscapes. The relevance of this aspect lies in the fact that existing decision rules used for computational agents may find only weak empirical support and thus require reevaluation and, where needed, readjustment in order to represent the reality of human behavior (Billinger et al., 2014).

Furthermore, there has been no experimental study that has investigated the search behavior of groups, either unstructured or structured. As discussed in the chapter on group related search, however, research has revealed that collectives of individuals tend to exhibit differing forms of behavior compared to individuals. Additionally, introducing structure or hierarchy to a group holds further potential for creating variation in observed behavior. The relevance of studying these aspects of behavior is grounded in the omnipresence of search tasks in the real world and especially in organizational contexts. Firms and businesses constantly face situations that can be described as requiring strategies for searching the problem space in order to find appropriate solutions. In that sense, a significant part of what businesses do can be captured using concepts related to the topic of search. Hence, deepening the understanding and expanding the knowledge on related issues can potentially benefit not only the academic community investigating such phenomena but also actors directly involved in situations affected by these.

This study aims at addressing the research gap that arises at the intersection of experimental studies on search and academic work related to the question of how groups perform in search tasks and how structure, in the form of an accountability mechanism, interplays in this relationship. As no experimental results exist for the search behavior of collectives of individuals and no investigation on the effects of structure in a group has been conducted, this work contributes to the understanding of how human searchers, as opposed to economic agents who follow stylized behavioral decision rules, solve problems presented to them that contain complex combinatorial decision structures and are by no means "solvable" by simply scanning the entire problem space for optimal solutions.

5 Effects of Structure

This chapter discusses important concepts and ideas from streams of academic literature dealing with structure in groups of people and the effects these can have on the behavior of the agents involved. As authority can be considered as one fundamental form of hierarchical order in collectives of human agents, and the ability to monitor and control is a central concept of authoritative relationships, accountability of individuals is a crucial idea used in this study to mimic this notion of responding to an authoritative entity. Subsequently, additional important concepts are presented and discussed.

5.1 Accountability

Accountability is a very popular concept in various different research fields and has also been used in the business literature (Cronshaw & Alexander, 1985; Peecher & Kleinmuntz, 1991). Accountability can be described as the implicit or explicit expectation that one may be asked to justify one's decisions, feelings or beliefs to one or multiple other parties (Lerner & Tetlock, 1999; Semin & Manstead, 1983; Tetlock, 1992; Scott & Lyman, 1968). It furthermore typically implies that depending on the quality of the feedback and how satisfactory the justification is, consequences are to be expected, ranging from social disdain and other punishments to respect and various kinds of rewards (Stenning, 1995). Accountability can be constituted in various forms and different types of relationships between agents have been used in the literature to depict accountability (Lerner & Tetlock, 1999). Among these are the mere presence of another, where people expect to be observed while performing a certain task (Zajonc, 1965; Zajonc & Sales, 1966; Guerin, 1993), identifiability of the agent, where individuals expect or know that their actions can be linked to them personally (Williams, Harkins & Latane, 1981; Zimbardo, 1969; Price, 1987), evaluation apprehension, where agents expect their actions and performance to be assessed and rated by someone (Guerin, 1989; Geen, 1991; Kimble & Rezabek, 1992; Simonson & Nowlis, 1996; Sanna, Turley & Mark, 1996), and reason-giving, in which case individuals expect to be asked to provide reasons and explanations on their motivation to undertake certain actions (Wilson & LaFleur, 1995; Simonson & Nowlis, 1998).

People who are observed while performing a task tend to try to comply to the audience's or the spectator's expectation, i.e. they will take into account the observer's beliefs on how and what should be done and try to conform to that. If, however, expectations are unknown or not clear, human agents will oftentimes engage in preemptive self-criticism which refers to the process of individuals thinking about their actions in more complex ways and trying to consider multiple perspectives on the issue at hand, including attempts to anticipate what others would do or think about their actions. In that sense, the mere feeling of being observed can lead to individuals engaging in more complex and cognitively more demanding ways of making up a decision (Tetlock, Skitka & Boettger, 1989; Mero & Motowidlo, 1995).

When individuals expect to be judged based solely on the outcome of their actions, regardless of the decision-making process leading to this decision, Siegel-Jacobs & Yates (1996) argue that this produces high levels of stress, as uncertainty about the success of the outcome is coupled with potentially negative decision consequences (Janis & Mann, 1977). This higher level of stress, compared to a setting where accountability is related to process rather than outcome, shows detrimental effects on decision makers behavior, reducing cognitive calibration and augmenting inconsistent behavior as agents try to increase their performance without knowing how to achieve this goal quickly. Note, however, that scholars of organizational behavior argue that outcome accountability is, in fact, beneficial for performance levels (Chubb & Moe, 1990) as it leaves decision makers with the freedom to achieve the goals set in any way they find appropriate. The argument is that private-sector institutions work more effectively, precisely because of this focus on outcome rather than process, whereas public-sector organizations put an emphasis on process by introducing bureaucratic regulation and enforcing conformity to these rules (Wilson, 1989). In that sense, this stream of literature poses a somewhat opposing notion to the effects of outcome accountability (Simonson & Staw, 1992; Lerner & Tetlock, 1999).

Predecisional accountability to an unknown spectator or audience can reduce bias if the bias stems from a lack of effort or self-critical awareness in the decision-making process. The mechanism behind this effect can be described as individuals who have to justify their actions, trying to avoid negative consequences, tend to engage in more thoughtful and critical behavior in order to avoid unfavorable feedback. This oftentimes results in people considering a wider range of alternatives and decision factors, anticipating different or opposing arguments to their own beliefs and overall approaching the task or decision with greater awareness of the cognitive process involved. For this change in behavior to have positive effects on performance, however, a lack of success in solving the task has to result from suboptimal levels of attention in the decision-making process and an improvement has to be possible without the need for additional knowledge or information, i.e. better performance is achievable through paying greater attention to the information available. If even after attention levels increase, no higher payoff can be achieved, accountability will likely have no effect on individuals' performance (Lerner & Tetlock, 1999). Lerner & Tetlock (1999) state that no amount of increased effort can compensate a lack of relevant knowledge when solving a problem. In that sense, overcoming a bias that produces suboptimal results, depends not only an individual's willingness to adjust their behavior but also on its ability to do so in terms of cognitive capabilities (Wilson & Brekke, 1996; Kerr, MacCoun & Kramer, 1996; Wegener & Petty, 1995).

Following these results, a frequent observation appears to be the fact that accountability only increases cognitive effort in certain situations and settings, while it is important to note that more cognitive effort must not always be beneficial for performance or the achieving of goals. The form of accountability revealed to be the most likely for causing cognitive efforts to be augmented, is predecisional accountability to an audience with unknown views, i.e. a situation where individuals expect to be observed by a spectator who doesn't prefer a specific outcome of the decision-making process but cares more about accuracy and how action is justified. Nonetheless, even with this very specific form of accountability, effects vary with regard to other factors. If a detrimental behavioral bias results from a lack of self-critical attention during the decision-making process, then holding decision makers accountable can improve performance, as long as such improvement is achievable without the necessity for additional knowledge or information. Under different circumstances, this form of accountability is likely to have no or even negative effects on existing biases or overall behavior (Lerner & Tetlock, 1999).

Accountability has also been found to enable individuals to adjust their decisions away from (suboptimal) anchors and, hence, improve the quality of their decisions (Kruglanski & Freund, 1983) and as Morrison & Milliken (2000) state in their work on organizational silence, there is strong evidence in related academic literature that people often feel threatened by negative feedback, be it about personal traits they possess or certain courses of action they approve of (e.g. Meyer & Starke, 1982; Swann & Read, 1981; Sachs, 1982; Carver, Antonio & Scheier, 1985). Consequently, human agents try to avoid receiving negative feedback (Ashford & Cummings, 1983). Vieider (2011) also shows that accountability increases cognitive effort among participants of an experimental study leading to those individuals partly overcoming well-known biases such as gain-loss framing situations and choices between simple and compound events. The author designs an experimental setup where effects of accountability and (monetary) incentives can be assessed separately and argues that these two, although oftentimes not differentiated in studies of incentives, produce different outcomes in terms of susceptibility to decision biases. His results are in line with numerous other findings in the accountability literature (Vieider, 2011).

5.2 Lack of control

An important aspect in cases where authority is involved in the relationships between agents is the concept of personal control of individuals. Personal control is defined as a person's belief in their ability to produce a desired change on the environment and individuals in organizations are typically considered to desire an increase in personal control. As this concept is based on beliefs, the term "control" does not necessarily reflect the actual ability of a person to cause "real" changes but can be understood as a subjective assessment of the person's situation in the organization. In that sense, the perception of control is the important aspect in this conceptual construct and it is this perception that can cause individuals to react if it doesn't meet desired levels (Greenberger & Strasser, 1986). When people feel a lack of control, i.e. the level of perceived control is below what is expected or what is regarded as the minimum requirement, this typically has detrimental effects on various aspects of an organization member's life within the hierarchical structure. Parker (1993) found that nurses who experienced a lack of (perceived) personal control in decision-making were more likely to quit and leave their workplace than colleagues reporting higher degrees of control. Those workers who felt high levels of control showed greater tendencies to attempt overcoming reasons for discontent as in "fixing" problems at the workplace rather than abandoning it altogether.

A frequently mentioned detriment to levels of perceived personal control is being watched or monitored while working on a problem. Aiello & Svec (1993) found that when people were monitored electronically while performing a complex task, performance achieved in this task decreased, whereas in other studies reducing a person's perceived personal control has been found to also decrease motivation and hamper cognitive performance (Seligman, 1975). The mere fact that one was being observed was reported by participants to cause stress and consequently exhibited a detrimental effect on the quality of work subjects were able to provide. Even without any consequences to be expected or any other kind of regulative or influencing measure, the act of monitoring a person's behavior turns out to generate enough stress and discontent that performance and other factors, such as satisfaction, for example, are impacted negatively. In other words, even having knowledge of the fact that one is being monitored, without information on the extent the monitoring is happening or the exact timing and implementation of it, can already act as a powerful enough stressor (Stanton & Barnes-Farrell, 1996).

5.3 Evaluation apprehension

Another well-studied concept related to the notion of being observed while performing a task is evaluation apprehension, i.e. the fear of being evaluated. The expectation of being evaluated has been shown to have effects on performance even if individuals are not expecting any consequences resulting from the assessment of their work. In fact, the expectation of evaluation is typically regarded to be strong enough to overcome social contexts and to impact productivity in a positive manner on its own (Sanna et al., 1996). This means that oftentimes it suffices for a person to think that an observer could possibly also be evaluating their performance, for evaluation apprehension effects to be present. Individuals who expect to be evaluated usually increase their productivity and effort when working on a task, in order to meet supposed expectations and avoid negative feedback.

Kimble & Rezabek (1992) conducted a study where they let participants play video games either alone or in front of an audience. The results showed that for complex tasks players always performed worse when being watched compared to when they were playing unobserved. This was true for both good players of the game as well as bad players. The phenomenon of individuals undercutting their own performance levels when exposed to spectators is also a prediction of the widely popular social facilitation theory which states that for simple tasks good performers will increase while bad performers will further decrease the quality of their results when having to perform in front of an audience (Zajonc, 1965). When performing a task in a group, individuals tend to not react as much to the prospect of being evaluated as in the case when they work on their own. The explanation appears rather intuitive: with other members of the group involved in the production of results, the responsibility for the outcome felt by individual group members presumably decreases. As now there is someone else to also be held accountable, the threat of evaluation shrinks. Note, however, that this can only be observed if subjects feel that they cannot be identified or that contributions to the final outcome cannot be assigned to single members of the group (Geen, 1991).

5.4 Responsibility attribution

Structure and the presence of authority also have an impact on people's perception of responsibility. Depending on the fact, namely, if, for example, a member of an organization has a superior to whom she responds to or is accountable to, the attribution of responsibility, influenced by the differing social context and associated roles, is likely to differ. Ways to assess an individual's responsibility in a given context, such as the focus on causing negative outcomes or the subject's overall level of obedience, are heavily influenced by the formation of social roles that form along the authority-subordination dimension (Hamilton, 1978). Central to the attribution of responsibility is furthermore a person's intention present in making decisions and taking actions, and the extent to which an outcome was produced by environmental forces outside the individual's control. In that sense, not only the actor's actual involvement in causing an outcome is relevant but situational factors play a role as well (Heider, 1958). Furthermore, in groups of individuals working together the phenomenon of diffusion of responsibility can be observed. This refers to individual members perceiving responsibility as something that can be divided, thus reducing one's own share. The fact that responsibility, and especially blame for undesired outcomes, can be "distributed" over all group members, is regarded as one of the factors that account for decision-making in collectives being typically more risk-seeking than individual behavior (Mynatt & Sherman, 1975).

5.5 Social comparison

It has been argued that people have an intrinsic desire of evaluating their abilities and opinions compared to other individuals (Festinger, 1954). This process of social comparison is present in a multitude of social settings and situations and is arguably a fundamental phenomenon in human cognitive behavior. Social comparison happens not only at the individual level within groups of actors but can also be observed between collectives. It oftentimes carries a certain notion of competitiveness where the own social identity is compared to others and desires to improve the own position arise (Turner, 1975). In general, three main types of motives that drive social comparison have been identified in the literature. The first one, evaluation, closely relates to Festinger's (1954) initial idea of social comparison being used as a mean for assessing one's one capabilities and other characteristics. In the original form, the focus lied on opinions and abilities, but subsequent research has constantly expanded the number of dimensions that can used for comparing oneself, resulting in the widely accepted view that virtually any aspect of an individual's being can be the basis for comparison. The second motive is about improvement and reflects the mentioned notion of using social comparison in order to improve own abilities. This aspect is argued to be closely linked to evaluation as one reason to learn more about oneself is the desire to be able to subsequently benefit from these insights. The third motive, enhancement, is defined as being specifically aimed at enhancing one's self-esteem or self-concept. In that sense, people engage in social comparison in order to feel better about themselves, justify certain actions or decisions or achieve similar goals related to the perception of oneself (Gibbons & Buunk, 1999).

6 Methodology

In this chapter, the chosen methodological approach is presented. First, the hypotheses derived from the discussion of the literature are developed, before addressing the experimental setup and variables used in the subsequent analysis.

6.1 Hypotheses

Based on the analysis and discussion of relevant academic literature and the concepts therein, four hypotheses are formulated in order to assess if any notable effects can be observed in search behavior when introducing groups, structured and unstructured. The notion of structure in this study relates to a very fundamental form of hierarchy: the presence of authority in the form of an accountability mechanism. This setting allows for analysis of effects of structure on a very basic level. As the authoritative entity has no real power to enforce consequences and is generally not more than an abstract presence, participants' behavior can be investigated in regard to this very "light" form of hierarchy, the idea behind that being to study effects of authority in a relatively "pure" version.

Before turning to structured collectives of searchers, unstructured groups in general are of interest. Although no experimental evidence on group search behavior in the context of NK-models is available, the literature provides a number of tendencies that can be expected from individuals performing jointly. Robust findings suggest that groups tend to be more risk-seeking, which can partly be ascribed to the diffusion of responsibility (Mynatt & Sherman, 1975), and Taylor & Greve (2006) found that team innovations produced greater variety in outcomes than individuals. These and related results, discussed in previous corresponding chapters, accordingly lead to the formulation of the first hypothesis:

H1: Groups of searchers tend to engage in more distant search than individuals.

Introducing structure in the form of an authoritative entity which holds individuals and groups accountable, the expectation on how this will affect search behavior revolves mainly around the notion that being watched or monitored causes stress and has been found to reduce performance due to a perceived lack of control (Stanton & Barnes-Farrell, 1996). Additionally, studies of evaluation apprehension have found that individuals perform worse in complex tasks when they are watched compared to when they act unobserved (Kimble & Rezabek, 1992). This is likely to create a tendency toward easily defendable decisions and, thus, the second hypothesis was derived as follows:

H2: Subjects will tend to engage in more local search under random justification

Similar to the lines of argument for this hypothesis, negative justification is expected to have comparable effects. What is important to add here is that due to this mechanism of producing negative feedback for every downturn in performance, subjects are expected to reduce explorative behavior even more. As this form of accountability clearly focuses on (negative) outcomes and thereby is likely to increase stress, the third hypothesis states that:

H3: Subjects under negative justification will tend to local search even stronger than under random justification

For the effects of the two accountability mechanisms on groups, a number of discussed findings are indicative. Results related to evaluation apprehension show that the related impacts of this concept are less severe for individuals in groups where single members cannot be identified (Geen, 1991). Social comparison theory suggests that when performing together the desire to evaluate one's ability and, furthermore, enable improvements thereby arises, bearing the potential for subjects to show increased effort and more explorative behavior when searching in a team (Gibbons & Buunk, 1999). Finally, the diffusion of responsibility typically observed in groups contributes to the notion of collectives being less receptive to the detrimental influences of accountability and additionally is associated with more risk-seeking behavior compared to individuals (Mynatt & Sherman, 1975). Hence, the fourth and final hypothesis is formulated as follows:

H4: Groups of searchers will be less receptive to both accountability treatments compared to individuals

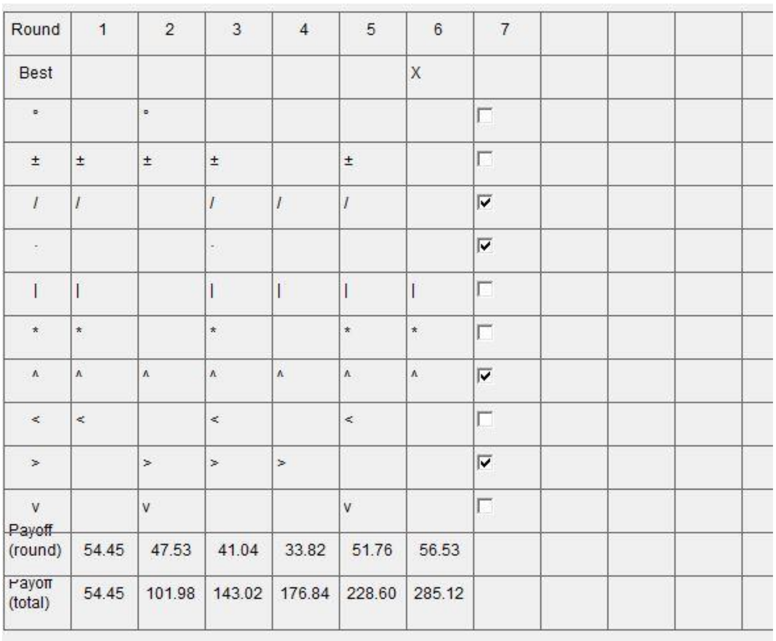
6.2 Participants

The experiment was conducted in the computer laboratories of the Vienna Center for Experimental Economics (VCEE) over the course of multiple days. Participants were recruited through the VCEE system and, in total, 160 subjects took part in the experiment, of which 56% were female. The age average was 25 and most subjects were students from the University of Vienna, although this was no requirement for participation. Subjects were compensated based on their performance in the experiment and earned an average of around 22 Euro for completing a 90-minute experimental session.

6.3 Experimental Setup

The experimental setup largely builds on the alien game as introduced in Billinger et al. (2014) and was programmed and conducted using the z-Tree software (Fischbacher, 2007). Subjects are presented a grid with ten graphical symbols, each symbol in one row, in the first column on the left-hand side. Every following column represents one of the 25 rounds that make up one trial or game. For every round, participants can choose a configuration of symbols that will be combined into a piece of art, subsequently sold to members of an extraterrestrial race to earn points. The idea behind this is that players have no other means of knowing which symbols or combinations will yield high payoffs other than engaging in trial and error, as no information is available on the preferences and likings of their customers. Each symbol can be toggled on and off by ticking the check box in the corresponding row. The number of symbols that can be combined is not limited in any way, meaning that players can create configurations consisting of none to all available symbols. Note that not selecting any symbol could be interpreted as submitting an empty canvas which would just as well lead to some payoff. In every round, players are free to (de-)select as many symbols as they wish. Therefore, one round's decision is not limited by any previous rounds' combinations or outcomes and every round starts with none of the symbols being selected, thus making this the default combination.

Additional to the choice variables, rows above and below the symbol grid carry further information. The first row, called "Round", displays the round numbers and indicates a player's progress in the current game. The second row is named "Best" and displays an "X"-symbol in the round where the highest payoff was achieved. This shall assist players in ensuring that they can easily identify their best performing configuration so far. Note that there is always only one “X” to be found in the row as even if the highest paying solution is repeated, the position marker remains at the round corresponding to when the most profitable configuration was discovered first. Furthermore, below the row for the tenth symbol, two more rows were installed, both related to the payoff players achieve. The upper one, named “Payoff (round)”, shows the payoff associated with the configuration of choice variables in the corresponding round, whereas the lower one, called “Payoff (total)”, displays the accumulated profit participants have earned up to every point in time during the game. Players are always presented with the entire grid spanning all 25 rounds, enabling them to analyze prior search trials and use all information gathered in the search process so far. Figure 1 below shows (part of) the input screen of the graphical interface of the game.



Round	1	2	3	4	5	6	7			
Best						X				
°		°					☐			
±	±	±	±		±		☐			
/	/		/	/	/		☑			
·			·				☑			
|	|		|	|	|	|	☐			
*	*		*		*	*	☐			
^	^	^	^	^	^	^	☑			
<	<		<		<		☐			
>		>	>	>			☑			
v		v			v		☐			
Payoff (round)	54.45	47.53	41.04	33.82	51.76	56.53				
Payoff (total)	54.45	101.98	143.02	176.84	228.60	285.12				

Figure 1: Input screen

The payoffs for every possible combination are drawn from one of two previously generated NK-landscapes with parameters N set to 10, for the ten graphical symbols acting as decision variables, and K fixed at 3, in order to ensure a sufficiently challenging but not excessively complex fitness landscape. The reason for using more than one landscape is that due to every subject playing the game twice, different payoff structures were needed for different conditions in order to exclude learning effects extending from one game or trial to the next. The two landscapes, however, have identical values for the N and K parameters and are similar in terms of minimum, maximum and average payoff as well as number of local peaks. Payoffs were calibrated to fall between 20 and 80 points for every possible combination of symbols.

6.4 Treatments

For this study, two kinds of treatments were defined that can be interpreted as “dimensions” spanning up a 3x2 matrix of six different conditions. The first treatment considers the lack of empirical and especially experimental work on how groups of individuals search compared to scenarios where search tasks are performed by single agents. Therefore, participants in the experiment played the game not only on their own but also as part of a 2-person collective, which represents the most fundamental form a team. By letting two players attempt the task together, it is possible to observe if and how collectives differ from individuals when it comes to search. In that sense, the differentiation of “Player Type” was used to classify conditions along this dimension, with “Individual” being considered the base case, whereas “Group” reflects the intended manipulation.

The second manipulation deals with the aspect of structured groups or organizations, especially the role of authority in those, and, hence, a mechanism of accountability is introduced. Specifically, participants face a situation where they might be asked to justify their decisions based on rules given by one of two specifications defined for this manipulation. In one version of the accountability treatment, from here on referred to as “Random Justification”, players face a certain probability of being asked for justification of a specific solution or submission after each round. The idea is to simulate accountability, and consequently a form of authority, by producing a setting where players feel that an external party is watching their performance and can request clarification for decisions made throughout the search process. The second variation of the treatment focuses more on the notion of being accountable towards a spectator that is clearly interested in a certain type of outcome. More precisely, this version named “Negative Justification” asks participants to provide a justification for their decisions only in cases when their performance decreases relative to their most recent submission. In other words, in every round where the payoff for the configuration submitted is lower than the profit from the previous round, players are asked to give an explanation on why they decided to choose this particular combination of symbols. The control group then is characterized by the absence of any of the two accountability mechanisms.

Every subject participated in two of the resulting six conditions, always playing once individually and once as part of a 2-person team, respectively. Furthermore, participants would either be part of the control group or of one of the justification treatments but would never take part in multiple accountability conditions. In other words, a participant would either (a) play the control version of the alien game on his or her own as well as with a second team member, (b) go through the random justification treatment, once individually and once as part of a team, or (c) complete the game with the negative justification scenario in both player type situations. The sequence of the individual and group conditions was varied over the course of the experiment so that some subjects would start on their own and afterwards form teams of two, whereas other participants would begin playing the game with a team member alongside and then proceed to rerun the same treatment individually. A table illustrating the six resulting variations constituting the structure of the experiment can be found below (Table 1). The following section gives an overview of the procedure of the experiment.

Player Type Accountability	Individual	Group
Random Justification	Individual – Random Justification	Group – Random Justification
Negative Justification	Individual – Negative Justification	Group – Negative Justification
Control / No Justification	Individual – Control	Group – Control

Table 1: Overview of conditions

6.4.1. Control condition

The control or no-justification condition represents the basic form of the game and is therefore used to explain the general procedure and main aspects of the experiment. At the beginning of every game, participants would be asked to enter their assigned identification number into the provided field and were then shown a short introductory text explaining what the task was and how the game works. After that, players had one trial round to familiarize themselves with the setup and make sure they understood the mechanics of the game. Every round is effectively split into four parts: first, the combination of symbols for this round is selected and submitted by clicking the “OK”-button in the lower right corner of the screen. Then, a short waiting period is implemented, followed by another screen that urges players to proceed by clicking “OK” again. Finally, the achieved payoff for the configuration submitted is presented. Note, that this structure is maintained for all conditions in the experiment with differences only in the content of the third step. Furthermore, it is important to mention that the payoff from the trial round was not added to the actual total payoff at the end of the game. As soon as the trial round was completed, subjects started the first of 25 rounds that made up one game. Players in the control condition run through the game unhindered and with no interruption in between rounds. Only the short artificial waiting period, determined randomly and lasting between three and seven seconds, was added between every submission and the corresponding payoff revelation in order to retain the necessary degree of similarity in regard to the accountability treatments. The motivation for implementing such a waiting time is provided in the following chapter.

The time limit for the two stages comprising one round, i.e. the configuration submission stage, where players select their desired combination of symbols, and the payoff revelation stage, where participants are presented the corresponding payoff of their submitted solution, was set to 30 seconds for each stage. In case players didn’t submit their configuration by manually clicking the “OK”-button before this time limit was reached, the submission was treated as if no symbols had been selected, resulting in the associated payoff for this round. Hence, it was crucial for searchers to confirm their choice in a timely manner in order to receive the correct profit. The importance of this aspect was mentioned in the introductory text and also stressed verbally in briefings before running an experimental session. As no interaction, except proceeding to the next round by clicking “OK”, was possible in the payoff revelation stage, no consequences were associated with exceeding the time limit in this case, other than the subsequent round or screen appearing automatically. The reason for implementing time limits was to keep durations of experimental sessions manageable. On average, one session, consisting of two separate games of 25 rounds each, lasted around 90 minutes, including briefings, reorganization due to player type changes and the completion of a post-experimental questionnaire.

The questionnaire was handed out after all subjects participating in the session had finished the game and consisted of questions related to demographic aspects of participants and a number of manipulation checks. As soon as the questionnaire was completed, subjects were able to collect their payments which were calculated based on the amount of points scored in total during both games of the session.

6.4.2. Random Justification

This treatment resembles the control condition in structural terms with the only exception being that after the artificial waiting period following every configuration submission, subjects face a chance of being asked to provide a justification for the particular solution they just submitted. As foreshadowed in the previous chapter, this waiting period was introduced to contribute to the credibility of the accountability mechanism. The basic idea was to evoke a sense of realism through inserting this "evaluation time" needed by the spectator before making a justification request. If such a request was to appear immediately after submission, it would likely not achieve the same impression of the submitted combination of choice variables being "under review". Note, however, that as no deception could be used in this experiment, participants in the random justification scenario were only given a minimal amount of information on that aspect of the game. More precisely, in the introductory text at the beginning of the game, players were informed that they may or may not be asked to justify their decision after a submission. No further information was given on why this happened, who was involved in this process or whether any consequences were to be expected. Furthermore, no additional explanation was provided as to how it would be determined if a configuration decision had to be justified or not, or how likely this was to happen throughout the game. In fact, the process of determining if an explanation, for why a certain combination of symbols was chosen, would be required or not, was a random one and the likelihood for a justification being requested was evaluated after every submission with the probability for that being fixed at 0.2 permanently.

If subjects were asked to elaborate on their motivation to decide for a specific configuration, a blank text field would appear on screen in which all comments had to be typed in. To avoid players skipping this step without providing any answer or just entering anything to continue, the minimum number of characters required to get to the next screen was set to ten. Even if the time limit was exceeded, there was no possibility of proceeding to the next screen without providing an answer meeting the minimum character requirement. The intention was to motivate people to actually engage in the justification of their decision and, furthermore, strengthen the credibility of the overall accountability mechanism. Subjects should feel as if their motivations matter to the spectator, which should contribute to a realistic setting regarding accountability and authority. Structurally, the justification stage was set after the waiting period and before the payoff revelation stage, replacing the screen simply urging subjects to continue via the "OK"-button used in the control condition. Note, that this screen is also used in the accountability treatments every time no justification request has been made.

6.4.3. Negative Justification

The negative justification condition differs from the random justification case in regard to the mechanism or rule used to determine if and when elaboration on decision motivations is requested from players. Where previously this aspect was determined randomly with a certain probability for each round, in this scenario the procedure focuses on the payoffs achieved by participants. Specifically, every time a submitted combination of symbols realizes a profit (strictly) lower than what was accomplished immediately before, i.e. if the payoff from the round prior to the current one is higher than what was achieved now, the submission is followed a justification request. This rule is also clearly stated in the introductory text at the beginning of each experimental trial and also displayed when the according justification screen appears. For an idea on how this screen looks like, refer to Figure 2 below (note, that the justification screen for the random justification condition is the same except for the indication that payoff has decreased).

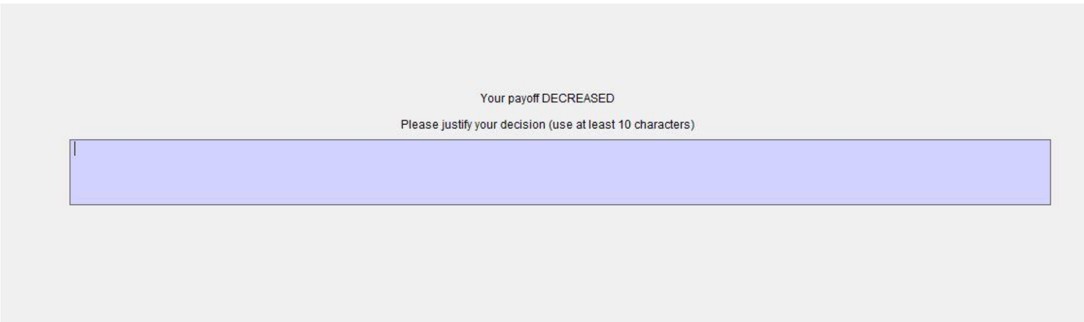


Figure 2: Justification screen

6.5 Player Type

Every variation on the accountability dimension was played both by individual players and by teams of subjects. After completing all 25 rounds of the first trial or game and being presented their final accumulated payoff for this part of the experiment, participants would, depending on the scheduling of their session, either proceed to form groups of two or replay the game on their own. As already mentioned, every session consisted of two games played by every subject taking part in this session. Both games were identical in terms of the accountability manipulation, so players would either face one of the justification treatments or the control condition of the game in both cases. The only difference, thus, being that one game was played individually and the other one as part of a group with a second player. Teams were composed of two subjects sitting next to each other in the computer lab and the initial assignment of seating was organized randomly. Members of a group would be seated together in front of one computer and would share one screen. Accordingly, the same version of the alien game was used as in the individual scenario and only the underlying NK-landscape describing the payoff distribution was changed to disable players from using knowledge from possible previous trials and, thus, to exclude strong learning effects. Communication within teams was not restricted in any way and no additional instructions or information was provided compared to the single player scenario.

6.6 Dependent variables

The focus of this work lies on the existence of potential differences in search behavior between individuals and groups in an interplay with authority in the form of accountability. In order to assess search behavior in general, it is important to define appropriate measures that reflect aspects of interest properly. One such aspect used throughout the academic literature on search is the scope of search employed by agents. As elaborated on in previous chapters of this work, the exploration-exploitation framework has been a popular tool for analyzing how search happens and how it differs under various circumstances. Following this approach and extending the corresponding line of thought, four different distance measures are used in the analysis of this study, three of which constitute new developments in the field.

6.6.1. Hamming distance

Among the most popular and relevant search measures in academic research, especially when using NK-models, is the Hamming search distance, introduced in Richard Hamming's (1950) work from the field of computer science. Used in numerous studies dealing with search behavior, including the original alien game study by Billinger et al. (2014), this measure captures the "closeness" of two solutions or locations on a fitness landscape in terms of how many single decision variables are configured differently in the two combinations. The higher the similarity between two configurations, the lower the corresponding distance between them, i.e. the closer they are located to each other.

Consider the following example to illustrate the construction of the Hamming distance: suppose a search problem with 3 decision variables where each of those can take either 1 or 0 as its value. Expressed in NK terminology, this means that N = 3 (the complexity parameter is not needed in this example). Assume that the current position is described by (000), so each of the three decision parameters shows the value 0. Moving away from this location to a point that shows a Hamming distance of one, relative to the current position, means changing any one, but only one, of the individual parameters to value 1. This implies that each of the locations (100), (010) and (001) can be described as being separated by a Hamming distance of value one from (000). Within a Hamming distance of value two, the following three locations can be found: (110), (101), and (011). Finally, the point (111) exhibits a (maximal) Hamming distance of three, relative to the initial starting point at (000). In simple terms, the Hamming distance is the sum of differences in decision variables between two configurations or solutions in the search space. The number of parameters that differ between two points on a fitness landscape directly corresponds to the associated Hamming distance. Changing only a few parameters in every new search trial and, thus, keeping the search distance low, can be referred to as narrow or local search, whereas performing larger reconfigurations in one attempt, resulting in high values for the Hamming distance, can be considered broad or distant search. In terms of the exploration-exploitation framework, local search translates to exploitative behavior, while distant search represents explorative tendencies (Billinger et al., 2014).

Following previous studies, the Hamming distance will, within the scope of this work, always be calculated relative to the best-performing configuration of choice variables that has been found so far. Assuming rational behavior and the desire to maximize the payoff throughout the game, it appears intuitive to consider the maximum-payoff configuration as some kind of "starting point" or, at least, reference point for future submissions (Billinger et al., 2014). Striving for high profits, subjects are expected to use their best solution as a guiding example for how they want to proceed. Hence, search distance, when measured using the Hamming distance, will always be stated relative to the best-performing solution. Measuring search distance relative to the most recent submission would also be a relatively straightforward approach but based on the literature and for the reasons mentioned this will not be considered for this study, in regard to the Hamming distance. One important aspect to note here is that this distance measure provides no information on the "direction" of where one point is relative to the other. In order to understand why this limitation can matter, especially in the context of this experiment, consider the following example: suppose an agent searches a landscape created by six binary decision variables and finds that the configuration (1100) provides the highest payoff discovered so far. Now suppose that, for whatever reason, the agent alternatingly chooses to submit configurations (0000) and (1111) for the rest of the search process. Considering Hamming distances evaluated after every submission, this agent shows a searching behavior of constantly selecting solutions that have a Hamming distance of two, relative to the best-performing configuration. By only taking into account this measure, it is impossible to differentiate this kind of behavior from someone who also only considers solutions within this distance but varies a lot more between them, e.g. selecting (1010), (0101) or any of the other combinations fulfilling the criterion of the Hamming distance being equal to two. Without additional information, these two, arguably very different, search behaviors would be treated as equal in terms of distance searched. Employing the exploration-exploitation framework, the two strategies would be classified as equally explorative although the latter one appears to clearly exhibit "stronger" explorative tendencies. An attempt to compensate for these shortcomings in analyzing search behavior led to the development of the remaining three distance measures used in this work.

6.6.2. Exploitation measures

How could one account for the "missing information" in Hamming distance measures that are responsible for the suboptimal classification of searchers in the example presented above? This question sparked the idea of developing a new form of measure to assess the search behavior of players of the alien game. A closer look on situations comparable to the mentioned example reveal that the central aspect for a new type of distance measure had to be the ability to account not only for the sum of choice variables differing from the reference configuration but to find a way to consider which parameters changed precisely. Put differently, it is not only important to know how much changed but also to know what exactly changed. Breaking this done to one number, comparable to the Hamming distance, was then the goal of the efforts leading to the present results.

Specifically, the additional analytical power was achieved by structuring the N-dimensional decision problem into sub-combinations. By forming tuples, i.e. groups of decision variables, consisting of at least two and up to N variables, the entire configuration is broken down to every possible sub-configuration and subsequently analyzed on that level. For both configurations of choice variables, for which the distance shall be calculated, all possible n-tuples are compared to one another and every time they are identical, a count variable is increased by 1. Thereby, all duplicate combinations, on all possible subsets of the complete solution, are counted and summed up. In that sense, the construction of this measure resembles, in principle, the idea behind the Hamming distance, with the extension that instead of only single variables being compared, this measure captures all possible subsets and composite solutions embodied in complete configurations. Although more demanding in terms of computational processing power than the Hamming distance, the basic working principle behind the production of this measure is therefore rather simple. The analytical and explanatory power, however, is potentially very promising.

Different types of n-tuples can be thought of as “tuple classes”, i.e. categories of tuples ranging from 2-tuples, 3-tuples all the way up to N-tuples depending the number of available decision variables. For each of these tuple classes the measure introduced can be calculated in the same way. As, however, absolute values of how many n-tuples have been identified to be shared by two configurations are of limited use to analysis, the sum for every tuple class is divided by the maximum possible number of duplicates. This maximum number of shared sub-combinations can be calculated by comparing two completely identical solutions, which, as they are duplicates of each other on the N-tuple level, consist entirely of duplicates on all other n-tuple levels as well. Alternatively, a simple way to calculate this maximum is to look at how many different n-tuples exist in a N-dimensional search problem in general. This normalization of the measure results in it always being a number between 0 and 1, with 0 indicating no duplicates, i.e. an entirely new or different solution, and, thus, maximally high exploration, and 1 representing an exact copy of the current position, i.e. an identical solution, which translates to the highest possible degree of exploitation. Additionally, this calibration allows to construct aggregate measures over all tuple classes by simply taking the average of values for all n-tuples from 2 to N.

The way how the measure is constructed also points to an important difference between the values of Hamming distances and those of the newly developed measure. Whereas high values for a Hamming distance indicate many differences between configurations and, thus, a large distance separating the two, results close to 1 for the introduced measure represent the exact opposite, namely, high similarity between solutions and, hence, great proximity in the search space. In that sense, this measure can be regarded as an exploitation, rather than an exploration, measure which exhibits high numerical values for increased levels of proximity instead of distance. Naturally, both the Hamming distance and this exploitation measure depict distances between solutions. The difference solely lies in the way how large and small numbers are assigned to the ends of this spectrum. Recalling the exploration-exploitation framework and the idea of these two concepts creating a continuum, the Hamming distance increases in numerical terms when exploration is present, whereas the new measure grows (numerically) with rising levels of exploitation. The importance of keeping this distinction in mind arises from the necessary caution that has to be taken when applying both measures simultaneously for data on search behavior.

The following example shall briefly recap and illustrate the construction of the exploitation measure: consider a simple search task with three choice variables (N = 3). Assuming the reference position to be (000) and a newly found solution to be compared as being given by (010), three 2-tuples can be identified for both configurations. The reference solution shows tuples [00], [00] and [00], whereas the newly found combination of decision variables can be divided into [01], [00] and [10]. For both cases, 2-tuples, the only available tuple class in this example, consist of the combination of choice one and choice two, choice one and choice three, and choice two and choice three, respectively. Comparing the two sets of tuples, it becomes apparent that only one duplicate exists, namely [00]. Note, that for the reference solution this tuple is not any arbitrary 2-tuple as all show the same values. It is specifically the second tuple, consisting of choice one and choice three, which is to be regarded as a duplicate compared to the newly found solution. The important point here is that it's not only about finding identical combinations in terms of values but about finding identical combinations of values that are in the same place. Hence, the number of identical sub-combinations is 1. Dividing this by the number of all existing tuples, which is 3, gives an exploitation measure for this example of 0.33, suggesting a modest level of exploitation in terms of 2-tuples.

It was mentioned in the previous chapter that three additional measures will be added while, so far, only the exploitation measure as a concept was introduced. The three measures are all exploitation measures, i.e. they are constructed and calculated in the way described above, and differ only in terms of reference configuration that is used for comparison. In the context of the alien game, this means that every round's configuration will be assessed in relation to three different reference points: the best solution discovered so far, the most recent submission, and all previous rounds' configurations. The first exploitation measure, relating to the best round and payoff so far, is the most similar to the Hamming distance as both share the same reference or comparison point. The second measure relates to the foreshadowed notion of using the last submission as a reference point being an intuitive approach and captures searchers' behavior in terms of fluctuation in search distance from one solution to the other. The third measure, finally, compares the current round's submission to all submitted configurations so far and attempts to put the focal solution in the context of the entire search behavior so far. Note, that this is the only measure that compares more than two rounds with each other. In all other cases, namely the Hamming distance, the last-best-round exploitation measure, and the last-round-exploitation measure, the current configuration is assessed in relation to one other round. Only for the all-previous-rounds exploitation measure are multiple rounds taken into account. As with all exploitation measures introduced here, this measure can only take values between zero and one, ensuring identical interpretation of results compared to the other two constructs.

7 Results

In this chapter, results of the regression analysis conducted on the data collected in the experiment will be presented. The following chapter then investigates hypotheses for validity and discusses implications of the findings.

For analysis of the experimental results a regression analysis was used. It tested for all four hypotheses and additionally takes a look at a number of other interesting aspects within the scope of the study. A regression for each of the four dependent variables, Hamming distance and the averages over all tuple classes for the three exploitation measures related to the prior round, the best round so far and all previous rounds, was conducted. Parameters used for the regression were the independent variables group identifier “Ind” and the treatments for random and negative justification as well as the control variables payoff achieved in the most recent round, i.e. the lagged payoff, an experience-related construct and the round number. The motivation behind using these controls is given below.

First of all, participants’ performance is in all likelihood linked to search behavior, i.e. the payoffs achieved by searchers are evaluated based on all four measures (dependent variables) for search behavior. This relates to a notion of performance-adaptive search where the reference configuration is the one used in the immediate predecessor round. Furthermore, the aspect termed “Experience” is introduced which captures the notion of subjects’ behavior changing depending on if they are playing the first or the second game in an experimental session. Although the payoff structures, i.e. the underlying NK-landscapes, are different for the two trials, a certain kind of learning effect, possibly linked to increased familiarity with the game in the second trial, cannot be excluded. In accordance with prior findings in the literature, search behavior is expected to adjust as the game proceeds, so the progress, represented by the round number a subject is currently in, was also investigated and linked to the search measures.

Additionally, two restrictions to the entire data set were implemented. The first one is related to the experience construct. As it became apparent that the fact that participants were playing the second trial influenced their search behavior compared to the first game, the decision was made to limit the sample to experienced searchers. The second limitation concerns the round numbers and states that only data from the second round shall be considered as none of the dependent variables can be constructed in the first round due to their comparative nature of calculation. These restrictions leave the regression analysis with 2,856 observations. Table 2 below shows the results of the regression analysis.

		Hamming	**Exp_allprev**	**Exp_last**	**Exp_lastbest**
Ind=1	coefficient	0.000	0.000	0.134	0.163
	std error	(.)	(.)	(0.08)	(0.08)
	p-value	.	.	0.099*	0.049**
Random=1	coefficient	-0.243	0.017	0.028	0.041
	std error	(0.23)	(0.06)	(0.06)	(0.06)
	p-value	0.287	0.776	0.655	0.468
Ind=1 # Random=1	coefficient	0.199	-0.027	-0.045	-0.060
	std error	(0.28)	(0.07)	(0.08)	(0.07)
	p-value	0.474	0.713	0.554	0.379
Negative=1	coefficient	0.464	-0.073	-0.066	-0.063
	std error	(0.23)	(0.06)	(0.06)	(0.06)
	p-value	0.042**	0.229	0.285	0.262
Ind=1 # Negative=1	coefficient	-0.444	0.001	-0.002	0.002
	std error	(0.28)	(0.07)	(0.08)	(0.07)
	p-value	0.112	0.986	0.977	0.982
L.Payoff	coefficient	-0.020	0.003	0.019	0.008
	std error	(0.01)	(0.00)	(0.00)	(0.00)
	p-value	0.001***	0.000***	0.000***	0.000***
Ind=1 # L.Payoff	coefficient	0.003	-0.002	-0.003	-0.003
	std error	(0.01)	(0.00)	(0.00)	(0.00)
	p-value	0.619	0.005***	0.006***	0.010***
Experience=1	coefficient	2.835	0.021	-0.768	-0.226
	std error	(0.35)	(0.05)	(0.07)	(0.07)
	p-value	0.000***	0.685	0.000***	0.001***
Ind=1 # Experience=1	coefficient	-0.091	0.078	0.000	0.000
	std error	(0.43)	(0.06)	(.)	(.)
	p-value	0.832	0.214	.	.
Round_No	coefficient	-0.048	0.011	0.016	0.024
	std error	(0.00)	(0.00)	(0.00)	(0.00)
	p-value	0.000***	0.000***	0.000***	0.000***
Constant	coefficient	0.000	0.000	0.000	0.000
	std error	(.)	(.)	(.)	(.)
	p-value	.	.	.	.
*** p<0.1, ** p<0.05, *** p<0.01**					
Observations	2,856				

Table 2: Regression Analysis

8 Discussion

This chapter first answers the question if the formulated hypotheses can be confirmed by discussing each of them sequentially. Findings are analyzed and discussed with regard to the implications they suggest. Thereupon, the remaining variables are assessed and light is shed on additional interesting insights the analysis provides.

8.1 Hypothesis 1

The first hypothesis stated that groups of searchers will engage in broader and more distant search than individuals. Results from the regression analysis confirm this at least partly. While no effects of collective search can be observed for the Hamming distance or the exploitation measure aimed at the entirety of rounds played so far, the exploitation measures related to the immediate prior round and the best-performing configuration discovered show weak ($p = 0.099$) and significant ($p = 0.049$) support, respectively. As the variable “Ind” was coded as 1 for individual players and the exploitation measures indicate non-exploratory behavior with increasing numerical values, the positive relationship between the two allows the reverse conclusion that groups indeed tend to search broader and engage in explorative efforts more often than individuals.

8.2 Hypothesis 2

For the second hypothesis no support could be found in the results at hand and, thus, this hypothesis has to be rejected. Possible explanations for the non-existence of variation in the search behavior with the introduction of random justifications could lie in the partly contradictory effects that accountability is likely to produce. Although significant parts of the related literature state that scenarios like this create stressful situations for individuals which render the occurrence of defensive, and therefore exploitative, behavior likely, other opposing tendencies might be responsible for cancelling these expected effects out. Arguments for this can be found in the observed tendency of human agents to engage in preemptive self-criticism when under accountability, which increases cognitive efforts for solving the task at hand (Lerner & Tetlock, 1999; Mero & Motowidlo, 1995). If exploitative behavior is derived from a person’s lack of effort during the search process, then such an increase in determination might likely be responsible for an opposing tendency arising which facilitates more explorative behavior and runs counter to the initially expected occurrence of more local and explorative strategies. A definitive answer to this question, however, requires additional work and probably more precise concepts and experimental designs that enable a separation of such counter-running trends.

8.3 Hypothesis 3

This hypothesis stated that negative justification will produce even stronger local search behavior among participants than the random justification treatment. The hypothesis has to be rejected as the results show not only a lack of more pronounced exploitative strategies but even reveal that there is a significant increase of explorative efforts under this form of accountability. While all three exploitation measures show no significant results, the Hamming distance grows significantly under negative justification ($p = 0.042$), suggesting that participants search farther when having to justify negative outcomes. This result comes somewhat surprisingly as apparently searchers' strategy for avoiding negative feedback is not sticking to known solutions and adjusting gradually but rather engage in explorative efforts to find superior configurations. This seemingly motivating effect of negative justification might be explained by an even stronger pressure to increase efforts due to the more stressful situation that this form of accountability was expected to create. Another important point here is that the results show no effects for individual searchers but only for groups, as the interaction between the variables "Ind" and "Negative" reveals. This means that only in collective search situations this rise in exploration can be observed. While single searchers apparently do not react to this form of negative feedback, groups clearly pursue a more distant search approach compared to situations of unhindered task solving.

8.4 Hypothesis 4

The expectation that groups react less to accountability mechanisms finds no support in the available data and corresponding results. While random justification was not able to produce any significant results, neither for individuals nor for groups, the negative justification treatment even caused reverse effects as groups were found to be the ones to react to this manipulation whereas single searchers showed no change in behavior. Accordingly, this hypothesis had to be rejected as well.

8.5 Other findings

Additional interesting findings concern the variables related to previous round payoff, experience in playing the game and progress in a trial. The first one, lagged payoff, shows highly significant results ($p < 0.01$) for all four measures of search behavior and reveals that as the achieved performance in the round before increases, searchers adapt to this feedback by narrowing down the search distance. In that sense, a high-performing solution leads to reduction in the Hamming distance observed and also causes an increase in all three exploitation measures used in this study. The reverse conclusion therefore states that disappointing configurations sparks more explorative behavior. This finding resonates strongly with the notion of adaptive search also observed in the study by Billinger et al. (2014) and, thus, can be considered in accordance with prior findings. Note, however, that there is one caveat to these results as individual searchers show no significant reaction to lagged payoff in terms of Hamming distance, while all exploitation measures still remain highly significant.

When it comes to searching with experience, i.e. playing the second trial of the game in an experimental session, participants show significantly higher levels of exploration and, accordingly, lower levels of exploitation. All dependent variables but one, the exploitation measure that captures the difference in regard to all previous rounds, reveal highly significant results ($p < 0.01$). Apparently, playing the game for the second time and, therefore, having built up a certain familiarity and confidence regarding how to play the game, leads to subjects engaging in riskier, i.e. more explorative, search behavior compared to the first trial. There is less exploitation of known solutions and the distances between new configurations and previous combinations of choice variables increases. However, there is again an important difference arising between single searchers and groups of players, namely, that only collectives reveal this tendency of more distant search after the first trial, whereas individuals do not seem to adapt their search strategies in a comparable way.

Finally, a player's progress in a search trial was revealed to have an impact on the search strategies employed. The variable carrying the current round number of a configuration shows a highly significant ($p < 0.01$) negative relationship with search distance. In other words, as the game advances, subjects tend to switch from explorative to more exploitative search behavior, meaning that they rely more on reaping the benefits of their search efforts by exploiting known solutions rather than pursue risky explorative strategies as the end of the game approaches. This observation is in line with previous theoretical and empirical work in the field and, hence, can be considered as confirming the corresponding expectation on how search behavior adjusts throughout the game.

9 Conclusion

The aim of this study was to address the question of how groups of individuals perform differently in search tasks compared to agents acting on their own. Existing literature on the topic of search on the one hand and group behavior on the other hand suggested that this might be a fruitful endeavor as a gap could be identified between these two realms. Scholars of group decision-making have, to date, not engaged in mentionable efforts to include search tasks in their empirical and, especially, experimental studies. Simultaneously, research on how individuals act when facing search problems was based heavily on the use of computational models and, thus, does not necessarily depict how human agents actually behave in such situations. Building on the recent work by Billinger et al. (2014), this gap was to be addressed by using an experiment where subjects would perform a search task as both individual actors as well as part of a group. Additional to introducing this element of collective search, the effects of structure, in the form of accountability, were also in the focus of this work. The expectation was that both structured as well as unstructured groups will show different search strategies compared to individual agents. The experimental setup used was largely based on the work by Billinger et al. (2014) and extended to address aspects of collective behavior and group structure.

In order to assess these questions, four hypotheses were formulated and evaluated based on the results of a regression analysis. All but the first hypothesis, however, had to be rejected. The second hypothesis, related to the expectation that random justification would produce more local search behavior due to increased stress, failed to be confirmed as this treatment showed no effects on the search strategies employed whatsoever. A possible explanation could be that there are other tendencies running counter to this trend, such as an increase in cognitive effort observed in other studies of accountability. The third hypothesis, expecting negative justification to cause even more narrow search among players, also found no support in the data, even revealing that this manipulation had a contrary effect for groups, resulting in collectives of players engaging in more explorative search when exposed to this form of accountability. Finally, the fourth hypothesis had to be rejected as it was formulated as the expectation of groups reacting less to the two accountability treatments. This expectation was not met as either there were no differences in search behavior between groups and individuals under the random justification treatment, or the negative justification manipulation even produced contrary results, revealing that collectives adjusted their strategies, whereas individuals didn't. For the first hypothesis, however, weak support could be found as groups did indeed engage in more distant search efforts compared to individuals on some of the dependent variables, i.e. search measures, used.

It is apparent, not only by looking at the results, that a lot of work has still to be done to arrive to a reliable level of knowledge and understanding concerning human search behavior under different conditions. The need for more data to ensure generalizability of experimental findings remains relevant irrespective of other aspects. The fact that the results of this study provided a number of surprising revelations is illustrative of the need for more research in the field. Numerous effects, counter-running tendencies and difficulties in separation and distinction between important concepts bear great potential in this mostly undiscovered field of search experiments. Finding new ways of modelling authority and hierarchy, developing new "games" and applications that can be used to illustrate search tasks and addressing existent concepts from related fields on a more detailed level are, thus, only some of the ways in which future research can continue the pursuit in this area. The relevance of enlarging efforts in this area also relates heavily to the fact that a majority of academic work on search behavior is based on simulations which, naturally, rely on assumptions and stylized facts derived from actual human behavior. To assess if these foundations are valid and find empirical support is crucial in validating findings produced by these kinds of scientific studies. Therefore, this work can be considered as taking a small step into that direction of providing insights into actual human behavior within the realm of search behavior and hopefully contributes to bringing more attention and research efforts to this fundamental and, hence, crucial part of human decision-making.

References

Aiello, J. R., & Svec, C. M. (1993). Computer Monitoring of Work Performance: Extending the Social Facilitation Framework to Electronic Presence 1. Journal of Applied Social Psychology, 23(7), 537-548.

Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. Organizational behavior and human performance, 32(3), 370-398.

Baumann, O., Schmidt, J., & Stieglitz, N. (2019). Effective Search in Rugged Performance Landscapes: A Review and Outlook. Journal of Management, 45(1), 285-318.

Becker, M. C., Knudsen, T., & March, J. G. (2006). Schumpeter, Winter, and the sources of novelty. Industrial and Corporate Change, 15(2), 353-371.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. Proceedings of the National Academy of Sciences, 115(35), 8734-8739.

Billinger, S., Stieglitz, N., & Schumacher, T. R. (2014). Search on rugged landscapes: An experimental study. Organization Science, 25(1), 93-108.

Bruderer, E., & Singh, J. V. (1996). Organizational evolution, learning, and selection: A genetic-algorithm-based model. Academy of management journal, 39(5), 1322-1349.

Burgelman, R. A. (1994). Fading memories: A process theory of strategic business exit in dynamic environments. Administrative Science Quarterly, 39(1).

Cardinal, L. B. (2001). Technological innovation in the pharmaceutical industry: The use of organizational control in managing research and development. Organization science, 12(1), 19-36.

Carver, C. S., Antoni, M. H., & Scheier, M. F. (1985). Self-consciousness and self-assessment. Journal of Personality and Social Psychology, 48(1), 117.

Chandler, A. (1962). Strategy and Structure: Chapters in the History of the American Industrial Enterprise (MIT Press, Cambridge, MA).

Chubb, J. E., & Moe, T. M. (1990). America's public schools: Choice is a panacea. The Brookings Review, 8(3), 4-12.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. The economic journal, 99(397), 569-596.

Cronshaw, S. F., & Alexander, R. A. (1985). One answer to the demand for accountability: Selection utility as an investment decision. Organizational Behavior and Human Decision Processes, 35(1), 102-118.

Cyert, R. M., & March, J. G. (1963). A behavioral theory of the firm. Englewood Cliffs, NJ, 2(4), 169-187.

Denrell, J., Fang, C., & Winter, S. G. (2003). The economics of strategic opportunity. Strategic Management Journal, 24(10), 977-990.

Doty, D. H., Glick, W. H., & Huber, G. P. (1993). Fit, equifinality, and organizational effectiveness: A test of two configurational theories. Academy of Management journal, 36(6), 1196-1250.

Festinger, L. (1954). A theory of social comparison processes. Human relations, 7(2), 117-140.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental economics, 10(2), 171-178.

Fleming, L. (2001). Recombinant uncertainty in technological search. Management Science, 47(1), 117-132.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. Research policy, 30(7), 1019-1039.

Floyd, S. W., & Lane, P. J. (2000). Strategizing throughout the organization: Managing role conflict in strategic renewal. Academy of management review, 25(1), 154-177.

Gavetti, G. (2005). Cognition and hierarchy: Rethinking the microfoundations of capabilities' development. Organization Science, 16(6), 599-617.

Gavetti, G., & Levinthal, D. (2000). Looking forward and looking backward: Cognitive and experiential search. Administrative science quarterly, 45(1), 113-137.

Geen, R. G. (1991). Social motivation. Annual review of psychology, 42(1), 377-399.

Ghemawat, P., & Levinthal, D. (2008). Choice interactions and business strategy. Management Science, 54(9), 1638-1651.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: development of a scale of social comparison orientation. Journal of personality and social psychology, 76(1), 129.

Gilfillan, S. (1935). Inventing the Ship (Follett Publishing Co., Chicago, IL).

Greenberger, D. B., & Strasser, S. (1986). Development and application of a model of personal control in organizations. Academy of Management Review, 11(1), 164-177.

Guerin, B. (1989). Reducing evaluation effects in mere presence. The Journal of social psychology, 129(2), 183-190.

Guerin, B. (1993). Social facilitation (Cambridge University Press, New York).

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The interplay between exploration and exploitation. Academy of management journal, 49(4), 693-706.

Hamilton, V. L. (1978). Who is responsible? Toward a social psychology of responsibility attribution. Social Psychology, 316-328.

Hamming, R. W. (1950). Error detecting and error correcting codes. The Bell system technical journal, 29(2), 147-160.

Heider, F. (1958). The Psychology of Interpersonal Relations (Wiley, New York).

Henderson, R. M., & Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing. Administrative science quarterly, 35(1), 9-30.

Jaikumar, R. (1986). Postindustrial manufacturing. Harvard Business Review, 64(6), 69-76.

Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment (Free Press, New York).

Jansen, J. J., Van Den Bosch, F. A., & Volberda, H. W. (2006). Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. Management science, 52(11), 1661-1674.

Johnson, D. S., & Garey, M. R. (1979). Computers and intractability: A guide to the theory of NP-completeness (WH Freeman, San Francisco).

Kabak, Ö., & Ervural, B. (2017). Multiple attribute group decision making: A generic conceptual framework and a classification scheme. Knowledge-Based Systems, 123, 13-30.

Kauffman, S. A. (1993). The origins of order: Self-organization and selection in evolution (Oxford University Press, New York).

Kauffman, S. A., & Weinberger, E. D. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. Journal of theoretical biology, 141(2), 211-245.

Kauffman, S., & Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. Journal of theoretical Biology, 128(1), 11-45.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. Psychological review, 103(4), 687.

Kimble, C. E., & Rezabek, J. S. (1992). Playing games before an audience: Social facilitation or choking. Social Behavior and Personality: an international journal, 20(2), 115-120.

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. Journal of experimental social psychology, 19(5), 448-468.

Laursen, K., & Foss, N. J. (2003). New human resource management practices, complementarities and the impact on innovation performance. Cambridge Journal of economics, 27(2), 243-263.

Lavie, D., & Rosenkopf, L. (2006). Balancing exploration and exploitation in alliance formation. Academy of management journal, 49(4), 797-818.

Lavie, D., Stettner, U., & Tushman, M. L. (2010). Exploration and exploitation within and across organizations. The Academy of Management Annals, 4(1), 109-155.

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. Administrative Science Quarterly, 52(4), 667-694.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. Psychological bulletin, 125(2), 255.

Levinthal, D. (1992). Surviving Schumpeterian environments: An evolutionary perspective. Industrial and Corporate Change, 1, 427-443.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. Management Science, 43(7), 934-950.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. Strategic management journal, 14(S2), 95-112.

Levinthal, D., & March, J. G. (1981). A model of adaptive organizational search. Journal of economic behavior & organization, 2(4), 307-333.

Levitt, B., & March, J. G. (1988). Organizational learning. Annual review of sociology, 14(1), 319-338.

Lippman, S. A., & McCall, J. J. (1976). The economics of job search: A survey. Economic inquiry, 14(2), 155-189.

MacDuffie, J. P. (1995). Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. ILR Review, 48(2), 197-221.

MacKenzie, D. (1992). Economic and sociological explanation of technical change. In Coombs, R., Saviotti, P., & Walsh, V. (eds.). Technological Change and Company Strategies: Economic and Sociological Perspectives. Harcourt Brace Jovanovich, London. 25-48.

March, J. G. (1988). Decisions and Organizations (Blackwell, New York).

March, J. G. (1991). Exploration and exploitation in organizational learning. Organization science, 2(1), 71-87.

March, J. G. (1994). Primer on decision making: How decisions happen. Simon and Schuster.

March, J. G., & Simon, H. A. (1958). Organizations (Wiley, New York).

Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. Proceedings of the National Academy of Sciences, 109(3), 764-769.

McGrath, R. G. (2001). Exploratory learning, innovative capacity, and managerial oversight. Academy of management journal, 44(1), 118-131.

Mero, N. P., & Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. Journal of Applied Psychology, 80(4), 517.

Meyer, W. U., & Starke, E. (1982). Own ability in relation to self-concept of ability: A field study of information-seeking. Personality and Social Psychology Bulletin, 8(3), 501-507.

Mihm, J., Loch, C. H., Wilkinson, D., & Huberman, B. A. (2010). Hierarchical structure and search in complex organizations. Management science, 56(5), 831-848.

Mihm, J., Loch, C., & Huchzermeier, A. (2003). Problem–solving oscillations in complex engineering projects. Management Science, 49(6), 733-750.

Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. American economic review, 80(3), 511-528.

Milgrom, P., & Roberts, J. (1995). Complementarities and fit strategy, structure, and organizational change in manufacturing. Journal of accounting and economics, 19(2-3), 179-208.

Miller, D., & Friesen, P. H. (1984). A longitudinal study of the corporate life cycle. Management science, 30(10), 1161-1183.

Mintzberg, H. (1979). The Structuring of Organizations (Prentice-Hall, Englewood Cliffs).

Morrison, E. W., & Milliken, F. J. (2000). Organizational silence: A barrier to change and development in a pluralistic world. Academy of Management review, 25(4), 706-725.

Mynatt, C., & Sherman, S. J. (1975). Responsibility attribution in groups and individuals: A direct test of the diffusion of responsibility hypothesis. Journal of Personality and Social Psychology, 32(6), 1111.

Nelson, R. R, & Winter, S. G. (1982). Evolutionary Theory of Economic Change (Harvard University Press, Cambridge, MA).

Page, S. E. (1996). Two measures of difficulty. Economic Theory, 8(2), 321-346.

Parker, L. E. (1993). When to fix it and when to leave: relationships among perceived control, self-efficacy, dissent, and exit. Journal of applied psychology, 78(6), 949.

Peecher, M. E., & Kleinmuntz, D. N. (1991). Experimental evidence on the effects of accountability on auditor judgments. Auditing: A Journal of Practice and Theory, 10, 108-113.

Perretti, F., & Negro, G. (2006). Filling empty seats: How status and organizational hierarchies affect exploration versus exploitation in team design. Academy of Management Journal, 49(4), 759-777.

Porter, M. E. (1991). Towards a dynamic theory of strategy. Strategic management journal, 12(S2), 95-117.

Porter, M. E. (1996). What is strategy?. Harvard Business Review, 74(6), 61-78.

Price, K. H. (1987). Decision responsibility, task responsibility, identifiability, and social loafing. Organizational behavior and human decision processes, 40(3), 330-345.

Puranam, P., Alexy, O., & Reitzig, M. (2014). What's “new” about new forms of organizing?. Academy of Management Review, 39(2), 162-180.

Rivkin, J. W. (2000). Imitation of complex strategies. Management Science, 46(6), 824-844.

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. Management Science, 49(3), 290-311.

Sachs, P. R. (1982). Avoidance of diagnostic information in self-evaluation of ability. Personality and Social Psychology Bulletin, 8(2), 242-246.

Sanna, L. J., Turley, K. J., & Mark, M. M. (1996). Expected evaluation, goals, and performance mood as input. Personality and Social Psychology Bulletin, 22(4), 323-335.

Schumpeter, J. A. (1939). Business Cycles (McGraw-Hill Book Company, Inc., New York).

Schumpeter, J.A. (1934). The Theory of Economic Development (Harvard University Press, Cambridge, MA).

Scott, M. B., & Lyman, S. M. (1968). Accounts. American sociological review, 46-62.

Seligman, M. E. (1975). Helplessness: On depression, development, and death. A series of books in psychology (WH Freeman/Times Books/Henry Holt & Co., New York).

Semin, G. R., & Manstead, A. S. (1983). The accountability of conduct: A social psychological analysis (Academic Press, New York).

Siegel-Jacobs, K., & Yates, J. F. (1996). Effects of procedural and outcome accountability on judgment quality. Organizational Behavior and Human Decision Processes, 65(1), 1-17.

Siggelkow, N. (2001). Change in the presence of fit: The rise, the fall, and the renaissance of Liz Claiborne. Academy of Management journal, 44(4), 838-857.

Siggelkow, N., & Levinthal, D. A. (2003). Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. Organization Science, 14(6), 650-669.

Siggelkow, N., & Rivkin, J. W. (2005). Speed and search: Designing organizations for turbulence and complexity. Organization Science, 16(2), 101-122.

Siggelkow, N., & Rivkin, J. W. (2006). When exploration backfires: Unintended consequences of multilevel organizational search. Academy of Management Journal, 49(4), 779-795.

Siggelkow, N., & Rivkin, J. W. (2009). Hiding the evidence of valid theories: How coupled search processes obscure performance differences among organizations. Administrative Science Quarterly, 54(4), 602-634.

Simon, H. A. (1955). A behavioral model of rational choice. The quarterly journal of economics, 69(1), 99-118.

Simon, H. A. (1957). Models of Man: Social and Rational (John Wiley & Sons, New York).

Simon, H. A. (1991). Bounded rationality and organizational learning. Organization science, 2(1), 125-134.

Simon, H.A. (1962). The architecture of complexity. Proc. Amer. Philos. Soc. 106(6):467–482.

Simonson, I., & Nowlis, S. (1996). Reasons and evaluations: The components of accountability in decision making. In JS Lerner & PE Tetlock (Chairs), Accountability: The social context of cognition. Symposium conducted at the 8th Annual Convention of the American Psychological Society, San Francisco, CA.

Simonson, I., & Nowlis, S. (1998). Constructive decision making in a social context: Unconventional choices based on reasons. Unpublished manuscript, Stanford University, 416-446.

Simonson, I., & Staw, B. M. (1992). Deescalation strategies: A comparison of techniques for reducing commitment to losing courses of action. Journal of applied psychology, 77(4), 419.

Smith, J. M. (1970). Natural selection and the concept of a protein space. Nature, 225(5232), 563.

Stanton, J. M., & Barnes-Farrell, J. L. (1996). Effects of electronic performance monitoring on personal control, task satisfaction, and task performance. Journal of Applied Psychology, 81(6), 738.

Stenning, P. C. (Ed.) (1995). Accountability for criminal justice (University of Toronto Press, Toronto, Canada).

Stuart, T. E., & Podolny, J. M. (1996). Local search and the evolution of technological capabilities. Strategic management journal, 17(S1), 21-38.

Swann, W. B., & Read, S. J. (1981). Acquiring self-knowledge: The search for feedback that fits. Journal of Personality and Social Psychology, 41(6), 1119.

Taylor, A., & Greve, H. R. (2006). Superman or the fantastic four? Knowledge combination and experience in innovative teams. Academy of Management Journal, 49(4), 723-740.

Teece, D. J. (1988). Technological change and the nature of the firm. In Dosi, G., Freeman, C., Nelson, R., Silverberg, G., & Soete, L. (eds.). Technical Change and Economic Theory. Pinter, London. 256-281.

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In Advances in experimental social psychology (Vol. 25, pp. 331-376). Academic Press.

Tetlock, P. E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. Journal of personality and social psychology, 57(4), 632.

Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. European journal of social psychology, 5(1), 1-34.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. Administrative science quarterly, 439-465.

Tushman, M. L., & O'Reilly III, C. A. (1996). Ambidextrous organizations: Managing evolutionary and revolutionary change. California management review, 38(4), 8-29.

Tushman, M. L., & Romanelli, E. (1985). Organizational evolution: A metamorphosis model of convergence and reorientation. Research in organizational behavior.

Tushman, M. L., & Rosenkopf, L. (1992). Organizational determinants of technological change: Toward a sociology of technological evolution. Research in Organizational Behavior, 14, 311-347.

Tushman, M., Smith, W., Wood, R., Westerman, G., & O'Reilly, C. (2003). Innovation streams and ambidextrous organizational designs: On building dynamic capabilities. Division of Research, Harvard Business School.

Usher, A. P. (1954). A History of Mechanical Invention (Cambridge, MA).

Vieider, F. M. (2011). Separating real incentives and accountability. Experimental Economics, 14(4), 507-518.

Vincenti, W. G. (1990). What engineers know and how they know it (Vol. 141). Baltimore: Johns Hopkins University Press.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: the role of naive theories in corrections for perceived bias. Journal of personality and social psychology, 68(1), 36.

Weinberger, E. D. (1991). Local properties of Kauffman's N-k model: A tunably rugged energy landscape. Physical Review A, 44(10), 6399.

Whittington, R., Pettigrew, A., Peck, S., Fenton, E., & Conyon, M. (1999). Change and complementarities in the new competitive landscape: A European panel study, 1992–1996. Organization science, 10(5), 583-600.

Williams, K., Harkins, S. G., & Latane, B. (1981). Identifiability as a deterrant to social loafing: Two cheering experiments. Journal of Personality and Social Psychology, 40(2), 303.

Wilson, J. Q. (1989). Bureaucracy: What government agencies do and why they do it (Basic Books, New York).

Wilson, T. D., & Brekke, N. (1996). Mental contamination and mental correction: unwanted influences on judgments and evaluations. Psychological Bulletin, 116(1), 117-142.

Wilson, T. D., & LaFleur, S. J. (1995). Knowing what you'll do: Effects of analyzing reasons on self-prediction. Journal of Personality and Social Psychology, 68(1), 21.

Winter, S. G., Cattani, G., & Dorsch, A. (2007). The value of moderate obsession: Insights from a new model of organizational search. Organization Science, 18(3), 403-419.

Wright, S. (1931). Evolution in Mendelian populations. Genetics, 16(2), 97.

Yassine, A., Joglekar, N., Braha, D., Eppinger, S., & Whitney, D. (2003). Information hiding in product development: the design churn effect. Research in Engineering Design, 14(3), 145-161.

Zajonc, R. B. (1965). Social facilitation. Science, 149(3681), 269-274.

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses. Journal of Experimental Social Psychology, 2(2), 160-168.

Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. Nebraska Symposium on Motivation, 17, 237-307.

Appendix

List of tables and figures

Abstract

Whenever people face a problem or have a task to perform, the need for finding a good solution arises. As numerous problems can be solved in various different ways, many of which potentially remain undiscovered for long periods of time, identifying the best solution is rarely a trivial task. The number of possible different solutions typically vastly exceeds the cognitive abilities of humans and frequently even machines. Constraints on resources such as time and processing capacities make an exhausting assessment of every thinkable solution to a problem practically impossible. In such situations, it becomes apparent why search is such a fundamental process in human decision-making. The discovery and evaluation of alternatives is a central aspect in countless activities ranging from simple everyday tasks to contexts of organizations taking decisions that affect a large number of people and can have far-reaching consequences. In the light of these examples, it becomes clear that studying human search behavior bears the potential for impactful discoveries and critical insights that can benefit decision-makers in a multitude of situations. Building on existing academic work in the realm of search behavior, this study aims at expanding current knowledge in the field by experimentally investigating differences in search strategies between individuals, unstructured groups and structured collectives.

Zusammenfassung

Immer wenn Menschen vor Problemen stehen oder eine Aufgabe lösen müssen, entsteht die Notwendigkeit eine gute Lösung zu finden. Da zahlreiche Probleme auf eine Vielzahl von Arten gelöst werden können, wovon viele potentiell für eine lange Zeit unentdeckt bleiben, ist das Identifizieren der besten Lösung selten eine triviale Aufgabe. Die Anzahl an möglichen Lösungen übersteigt typischerweise die kognitiven Fähigkeiten von Menschen und häufig sogar jene von Maschinen. Einschränkungen bei Ressourcen wie Zeit und Rechenleistung sind oft dafür verantwortlich, dass eine erschöpfende Bewertung aller denkbaren Lösungen praktisch unmöglich ist. In solchen Situationen wird ersichtlich, warum Suche einen fundamentalen Prozess der menschlichen Entscheidungsfindung darstellt. Die Entdeckung und Bewertung von Alternativen ist ein zentraler Aspekt in zahllosen Aktivitäten, von einfachen Aufgaben im Alltag bis hin zu Fällen, in denen Organisationen Entscheidungen fällen, die eine große Anzahl an Menschen betreffen und weitreichende Konsequenzen haben können. Im Lichte dieser Beispiele wird klar, dass eine Beschäftigung mit menschlichem Suchverhalten einflussreiche Entdeckungen und entscheidende Einblicke liefern kann, von denen wiederum Entscheidungsträger in einer Vielzahl von Situationen profitieren können. Aufbauend auf bestehenden akademischen Arbeiten auf dem Gebiet des Suchverhaltens zielt diese Arbeit darauf ab die gegenwärtige Wissensbasis in diesem Bereich durch experimentelle Untersuchungen zu Unterschieden im Hinblick auf Suchstrategien zwischen Individuen, unstrukturierten Gruppen und strukturierten Kollektiven zu erweitern.