# ADVICE UTILIZATION AND DECISION PERFORMANCE UNDER RANDOM ADVICE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ALİ ÖZARSLAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF BUSINESS ADMINISTRATION

JUNE 2014

Approval of the Graduate School of Social Sciences

_____

Prof. Dr. Meliha Altunışık
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in the Department of Business Administration.

_____

Prof. Dr. Cengiz Yılmaz
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master Science in the Department of Business Administration.

_____

Assoc. Prof. Dr. M. Sinan Gönül
Supervisor

**Examining Committee Members**

Prof. Dr. Özlem Yılmaz (METU, BA) _____

Assoc. Prof. Dr. M. Sinan Gönül (METU, BA) _____

Asst. Prof. Dr. Ayşe Kocabıyıkoğlu (Bilkent, MAN) _____

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last Name: Ali Özarslan

Signature :

# ABSTRACT

ADVICE UTILIZATION AND DECISION PERFORMANCE UNDER RANDOM ADVICE

Özarslan, Ali

M.B.A., Department of Business Administration

Supervisor: Assoc. Prof. Dr. M. Sinan Gönül

June 2014, 95 pages

The importance of advice taking in decision making has been arising for two and a half decades. The existing research in advice taking has suggested numerous measurement techniques that are utilized under different types of advice such as random advice and perfect advice. In this thesis using the random advice framework, through an experimental design, I compare some of formula-based, regression-based and scale-based techniques of advice utilization that are widely used in judgment tasks and test the decision performance of decision makers based on advice utilization. Results and general findings derived from these experiments are summarized and suggestions for future research on advice taking are offered.

Keywords*:* Random advice, Advice taking, Judge-advisor system, Advice utilization, Decision analysis.

# ÖZ

RASSAL TAVSİYEDEN YARARLANMA VE KARAR PERFORMANSI

Özarslan, Ali

Yüksek Lisans, İşletme Bölümü

Tez Yöneticisi: Doç. Dr. M. Sinan Gönül

Haziran 2014, 95 sayfa

Karar alırken tavsiye almanın önemi son 25 yıldır gitgide artmaktadır. Öneri alma konusundaki mevcut çalışmalar tavsiye kullanımıyla ilgili rasgele veya doğru tavsiye gibi farklı tavsiye türlerini göz önünde bulundurarak çeşitli numerik teknikler kullanmışlardır. Bu tezde rasgele tavsiye metodu kulanılmış ve deney çalışması yoluyla yargısal tahminlerde çokça kullanılan numerik teknikler (formül-tabanlı, regresyon-tabanlı ve ölçek-tabanlı) karşılaştırılmış ve insanların öneriden yararlanmasıyla kararlarlarındaki tutarlılık performansı arasındaki ilişki incelenmiştir. Tespitler ve sonuçlar açıklanarak ileriki dönemler için hangi alanlar üzerinde çalışılabileceği önerilmiştir.

Anahtar Kelimeler: Rassal öneri, Öneri alma, hakem-danışman sistemi, Öneri kullanma, Karar analizi.

*To my lovely wife Derya who is the source of my success with her unconditional support and love.*

*To my parents who always encouraged me in any moment of crisis.*

*To the people who have spent their valuable time for me.*

# ACKNOWLEDGMENTS

First of all, I would like to express sincere gratitude to my supervisor Associate Professor M. Sinan Gönül who has provided very valuable guidance and encouragement during my thesis. As you devoted a lot of time to guide me throughout my thesis, your humble supervision and support enabled me complete my thesis without any problem.

I am also grateful to my thesis committee members, Prof. Özlem Yılmaz and Asst. Prof. Ayşe Kocabıyıkoğlu for their valuable comments and suggestions for the further study.

My appreciation also goes to many friends at METU and Bilkent University for their contribution and support.

# TABLE OF CONTENTS

# LIST OF TABLES

TABLES

# LIST OF FIGURES

FIGURES

# CHAPTER 1

# INTRODUCTION

## 1.1 Decision Making

People always have difficulty in dealing with uncertainty while making typical decisions which may be as simple as following the appropriate path to reach a place, or more complex as choosing the right university to study. It is an intrinsic need to resolve the uncertainty in judgments (Festinger, 1954), however arduous this task is. Fortunately, many times, such real life decisions are facilitated by incorporating help from others. That is partly because people prefer to benefit from additional opinions in their decisions. Indeed, research of Solomon Asch (1952) showed that individuals choose to agree with group's judgment even if it is apparently wrong. Furthermore, studies show that people's perceptions and choices are seriously influenced by others' opinions (Ajzen & Fishbein, 1980). In short, people are ready to be influenced by others and change their opinions to some extent.

People always want to make more accurate judgments and decisions in any situation. In order to improve judgment accuracy, people should efficiently use additional opinions in their decision. This ideal leads to the question "How do people utilize help from different sources while making decisions?". For instance, imagine that you are coach of a football team that starts the second half of the game in a defeated position. Your assistants and you have different tactics for the rest of the game. Whether assistants' recommendation will be effective on coach's initial opinion and which facts

will play role on the final decision of coach are parallel to the question above. The influence of advice on final decision of people is analyzed within advice giving and taking literature.

In this context, opinions of others and how people perceive them are the critical components of advice taking in decision making. Receiving advice while the decision maker (hereafter judge) also holds his/her own opinion leads to an intrinsic conflict in reaching the final decisions. It is the decision maker's responsibility to combine multiple sources of information and reach the final decision. The decision making process of a judge moderated by the advice is examined in terms of several aspects including the complexity of decisions, advisor confidence, number of advices and the interaction between judge and the advisor (Bonaccio & Dalal, 2006). Before mentioning those factors, making judgments under uncertainty will be briefly introduced and then cognitive, social and psychological aspects of advice taking are explored.

### 1.2 Judgment under Uncertainty

In a typical textbook problem, there exist questions asking the exact answer of the problem and also some data is given that helps reaching the solution. However, real life problems do not come as a complete textbook problem. There is always an ambiguity that affects the problem and thus, judgment plays a huge role in the decision making under the case of such uncertainty. The likelihood of an uncertain event can be expressed by assessing probabilities to the sub-events that are influencing the major event. For instance, the close of Dollar/TL on Monday depends on the decisions of FED during the day, but it also depends on the Dollar sale or purchase of central

bank of Turkey. Consider that only these two events affect the Dollar/TL value, so assigning probabilities to those two events and also predicting the effects on the exchange rate requires deep involvement of judgment.

Most decisions are structured in a complex manner. The complexity comes from the variety of objectives, options, parties as well as the uncertainty in most of the cases. Moreover, human mind has limited resources and capabilities. The decisions become even more difficult when the time constraint is included. For these reasons, people always get in trouble with complexity and tend to find simple ways to solve problems. Specifically, judgmental operations on an issue with limited data enforce the decision maker to hold on some heuristic principles to reduce the complexity of the issue (Tversky & Kahneman, 1974).

According to Simon (1982), people utilize some approximation methods so as to deal with complex decision problems. These methods are called heuristics and they are involved within cognitive processes used to make decisions in a shorter period. In other words, by using the laws of inference, people try to find cues that help to reach to reasoning for the decision preference (Hardman, 2009). Herbert Simon devised the "bounded rationality" concept and dedicated it to the limitations of human mind and environmental structure (1981). He claimed that one form of bounded rationality is satisficing that is choosing the alternative that meets the criterion instead of looking for the optimal one. However, satisficing is not a preferable method for the ones that always seek for optimality. Thus, bounded rationality hurts the idealistic decision process.

Tversky and Kahneman (1974) explored three heuristics that decision makers use. They asserted that reliance on judgmental heuristics result in cognitive biases. Nevertheless, they referred those heuristics as economical and effective in decision process. Initially, any event or situation that can be represented by well-known facts or stereotypes is easy to be examined. By this way, the degree to which an event is representative of another one is used in exploring the probabilities assigned for that. Tversky and Kahneman (1974) called it "representativeness heuristic" which arrives from the nature of the connection between events or facts.

The second heuristic, "availability" occurs when people retrieve, remember and consider the occurrences of events that they face or imagine and then decide based on these memories. Another heuristic is called anchoring which arises due to the existence of initial estimates given to the decision maker. The initial values may be set as extreme values, but people are always affected by this information giving some credibility to this as if it is someone else's opinion and so they change their opinion. Deviating from initial opinion results in biases and this causes final estimates that are different from the decision maker's original opinion.

A more benevolent perspective on these mental shortcuts is presumed by Gigerenzer and his colleagues (1999). Kahneman, Slovic and Tversky (1982) called them quick and dirty heuristics whereas Gigerenzer and Todd (1999) named those heuristics as "fast and frugal". Since these inferences help decision makers to decrease their effort of time and knowledge, they are fast and frugal. For example, recognition heuristic can be used in environments where an alternative is more recognized than another one. Moreover, minimalist, take the last or take the best strategies work well in different

circumstances, but a decision maker reaches to a single cue by using these heuristics and then decide based on that cue. This is not a preferred way of decision since optimality concern is mostly desired.

To sum up, uncertainty and complexity of decision problems lead people to engage in judgmental activities, but individual judgment requires relying on cognitive heuristics. The heuristics summarized above provide quick paths for reaching the final decision. However, those heuristics create some conflicts and errors which may cause misleading judgments and inaccurate decisions (Humphrey, Hollenbeck, Meyer and Ilgen, 2002). In addition to that, they are not the quick ways to reach the optimal decision (Gigerenzer and Todd, 1996). One way to decrease the possibility of an outcome such as non-optimal decision is to use additional opinions before making decisions. Supporting this view, it is necessary to investigate the literature on group and team decision making. Some of the main findings of this literature will be explored in chapter 2.

## 1.3 Organization of the Thesis

People always encounter with judgmental activities during deciding on complex issues and individual judgments suffer from cognitive processes exposed by the limited human mind. In order to increase the accuracy on decisions, people revise their judgments by getting advice from others. This is true in organizational settings as well, since managers as leaders mostly consult their co-workers or employees while giving decisions about company. Since the use of advice in real life decisions is apparently dominant, scholars suggested examining the advice giving and taking framework as Yaniv (2004b) asserted "It is imperative for future research to

consider the procedures by which various type of advice are elicited and used best" (p. 12).

In this thesis, I examine to what extent people utilize advice and whether utilized advice increases the decision performance. In doing so, I prefer to use random advice that is generated by a software with a bounded error. In other words, the advice is selected at random from a normal distribution centered on the true estimate with a constant standard deviation. There are lots of studies that have utilized random advice in their experimental designs (e.g., Fischer & Harvey, 1999; Clarebout & Elen, 2008). Cohen, Oates and Amant (1996) found that routing the cargos by utilizing random advice is as good as the case of an agent's advice. It can be inferred from previous studies that it is convenient to use random advice as it does not result in any negative outcomes.

The main aspect behind the usage of random advice is that random advice can represent both expert and non-expert advice. The argument is that people may not assess the quality of every advice in professional or daily life. An advice might be good or bad advice depending on the expertise and knowledge of the advisor. Nevertheless, they cannot be perfectly sure that an advice will help to increase decision performance even if it comes from an expert. Thus, that advice might be simulated by a random advice as it can reflect good, moderate or bad advice. Studies that use random advice prefer it due to its appropriateness and statistical power. First of all, it is easy and time saving to use random advice and also it can be generated in form of expert or novice advice depending on the situation by controlling the error term in the analysis. Secondly, randomization in any field leads to statistical conformity due to the fairness and lack of bias. Statistical power of

randomization leads scholars to utilize randomization in their analysis (e.g., Ferron & Sentovich). For these reasons, random advice is generated by controlling the error term.

After forming the random advice for the questionnaire, experiment is done through multiple decisions. Here, I follow a unique and novel approach and categorize the decisions with respect to utilized or discounted advice. In other words, when people are introduced with an advice in a problem, they may totally change their ideas in a positive or negative way. This leads to three situations in which initial opinion, advice and final opinion take different ordering except in the case of equalities (e.g., advice (A) >initial estimate (IE) >final estimate (FE)). For the six situations, a specific item falls between the other two items in two cases (e.g., IE<A<FE and FE<A<IE). By combining these two cases of three situations, the decisions are categorized into three types.

After dividing the decisions into three parts, the three orderings are named as extreme adjustment, bracketing and moderate adjustment. Here moderate adjustment is the usual case in which judges use both advice and their pre-advice estimates whereas the other two situations are rarely observed since people mostly solicit opinions of an expert. Extreme adjustment is judges' over tendency to move far away from initial decision and estimate in favor of an advice. Bracketing term is first used by Soll and Larrick (2009). The authors combined the estimates of two judges by assuming that one of them is a judge and the other is an advisor. When estimates of two judges (let them be 70 and 100) fall in the opposite side of true estimate (90), that estimates bracket the truth. In present case, final estimate and advice bracket the initial estimate.

The reason of categorizing the adjustment behavior of the decision makers is that firstly categorizing with respect to adjustment helps to analyze the usual case (moderate adjustment) with unproblematic data since case that advice is unutilized is separated from it. On the other hand, previous studies do not focus on the rare situations. It is suggested by Bonaccio and Dalal (2006) to study these extreme conditions. I wonder the utilization rate of the extreme decisions and whether the accuracy is improved or not in those decisions. The decision performance of unusually behaving judges may give some evidence about the reasons of such behaviors. For instance, it is expected from bracketers to have high confidence considering that more confident judges solicit less advice as Cooper (1991) supported. In short, an important contribution of this thesis is the analysis of the extreme situations in advice taking by carrying out different measurement techniques.

After categorizing the decisions with respect to adjustment ordering, firstly I compare formula-based, regression-based and scale-based advice utilization measurement techniques on three different adjustment styles. These techniques are used in judgment tasks that involve quantitative decisions such as estimating probability, score or date of certain events. Then I test the decision accuracy of judges considering whether utilized or discounted advice lead to better results in decisions. In doing so, I initially compare the decision performance of the usual and extreme cases. Then some methods are used to generate estimates that are used as a benchmark in comparison of the performance of the judges that utilized or discounted advice.

The rest of the thesis is structured as follows: Chapter 2 starts with the individual and team decision making frameworks. It is revealed that advice taking framework contributes to understand the dynamics of team decision

making. Then advice and advice utilization concepts are defined and motives behind advice taking are explained. In the following parts, judge-advisor system (JAS) is introduced and factors affecting advice utilization and decision performance in JAS framework are described through literature review of advice taking and utilization. Lastly, some techniques on measurement of advice utilization (formula-based, regression-based and direct rating approach) are mentioned in a detailed manner.

After description of advice taking literature, in chapter 3, the design and methodology of the study is introduced. Firstly, categories of adjustment are introduced. The three categories will be used throughout the thesis. Then procedures of study are explained in a detailed manner. In chapter 4, results of formula-based, regression-based and direct rating approaches are explored. Then results of these methods are compared and decision performances of the three cases are described. Lastly in chapter 5, discussions are made regarding to the methodology and results of the study and suggestions for further research are given.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Individual vs. Team Decision Making

### 2.1.1 Individual Decision Process

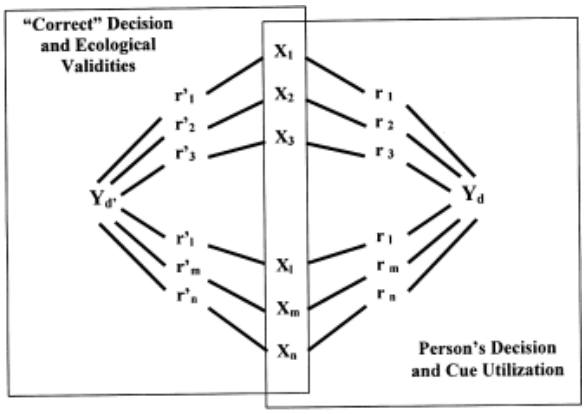


**Figure 1 - Brunswick's (1948) Lens Model of Decision Making (Adapted from "Multilevel Theory of Team Decision Making: Decision Performance in Teams Incorporating Distributed Expertise" by J. R. Hollenbeck, D. R. Ilgen, D. J. Sego, J. Hedlund, D. A. Major, J. Phillips, 1995, Journal of Applied Psychology, 80, p. 295.**

One of the oldest decision making model developed by Brunswick (1955, 1956) is called as the lens model. Brunswick used linear regression in his model by assigning probabilities to each cue that helps the decision maker to reach to the decision. In lens model, set of cues are evaluated by using functional relationships in order to fit to the decision in the best way. The left-hand portion of the model shown in figure 1 is the correct model with actual criterion and optimal weighting scheme. Ecological validity is related to the relationship between cues and criterion used in a real world situation.

Exploring all relevant cues, assigning weight to each cue and reaching to a decision are the steps of optimal decision making. However, this is not the case for most real life situations. Human judgments suffer from limited capabilities of human mind as well as the biases resulting from heuristics. Thus, cues are not perfectly utilized leading to a gap between a person's decision and the optimal one (March & Simon, 1958). Brunswick's lens model summarizes the individual decision process by presenting the ideal framework, but the practical implications of the model are limited in organizational settings, because much work is done through small groups.

### 2.1.2 Team Decision Making

Teams are generated by a group of people in order to complete tasks or solve problems by giving decisions. In a formal team decision making process, the group acts as a basic entity and the decisions are given by the group. Group members give some pieces of advice and those pieces of advice are judged by the group or the leader. In organizational context, teams are generated by top managers and common goals are set in any team during the forming stage. Even top management is an example of team in which important decisions about an organization are made. The importance of team decision making led the scholars to give particular attention to the decision processes of team settings (Ilgen, Major, Hollenbeck & Sego, 1995). Studies revealed that dynamics of team settings play an important role on the decision performance (e.g., Sundstrom, De Meuse & Futrell, 1990).

There are different types of team settings depending on the existence of hierarchy. Jury type teams are one of the typical examples for consensus-based groups. Since juries are generated by individuals with different

expertise and no formal leader exists in a jury, selection process turns to a paradigm about how to reach final decision. That is why consensus-based group setting is not preferred in decision making tasks. On the other hand, status difference in such teams ensures that the team has a leader who is the actual decision maker preventing the team from conflicts and misjudgments. Hollenbeck and his fellows (1995) called it "hierarchical teams" in which one member, the leader is responsible for the decision and each member contributes to the team by own knowledge and judgment.

Vroom and Yetton's contingency model is a framework that is based on the leader's situation in an organization. According to the model, there are five hierarchical forms of decision making depending on the decision making mechanism (1973). The first style is autocratic approach in which the leader does not consult for additional opinion and gives the decision based on the available information. The most social form is the group-based style in which the team decides on consensus basis by agreement or voting. The other three styles involve both the individual opinions and the leader's responsibility for decision. It is revealed that participative decision making is getting dominant in organizations. The model indicates that hierarchical decision making is required in case of quality consideration and insufficient information. Vroom-Yetton model suggests that the leader should choose the appropriate form of decision making depending on the dynamics of situations. However, the model is criticized in terms of the rationality of the leader, because bounded rationality theory indicates the biases that result from the limitations of human mind and environmental structure.

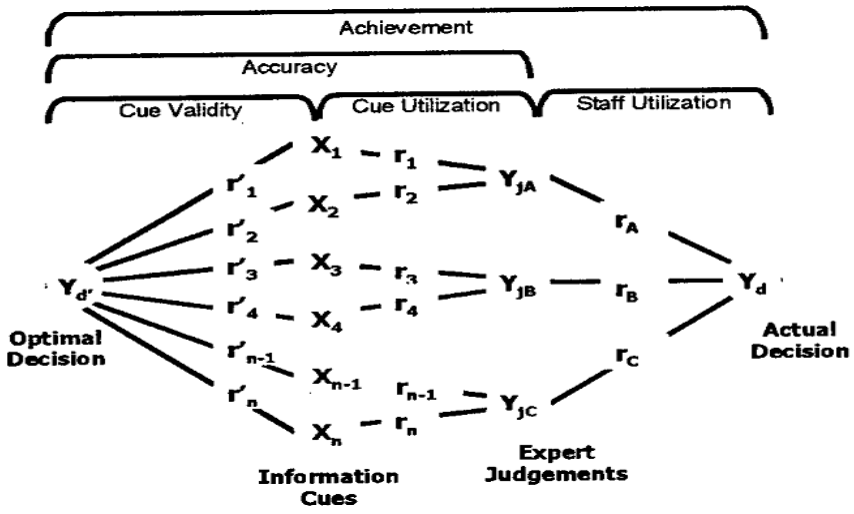


**Figure 2 - Brehmer and Hagafors's (1986) Model of Staff Decision Making (Adapted from "Using the Same Decision making Process for Joint and Army Operations" by K. R. Smith, 1999, School of Advanced Military Studies, p. 10)**

Brunswick's lens model expresses the ideal decision making process, but it does not involve any team structure. As shown in figure 2, Brehmer and Hogafors (1986) added the hierarchical team setting to lens model and called it "staff decision making model". In this model, each member is assigned with different cues and gives individual opinion on final decision. The leader of the team combines those recommendations, analyzes and then gives the decision. Staff members utilize from information cues by giving weights and the leader judges the weights of each member by comparing cue utilization and validity. The authors explored that low accuracy can be due to the low cue validity and utilization. In both cases, the leader highly utilizes its staff although he or she has no idea about the information cues.

Ilgen and his fellows (1995) contributed to the model by considering the leader's initial opinion about cues. Moreover, they pointed out that cues can be relevant to each other, so team members need to communicate during the decision process. Hollenbeck and his fellows (1995) developed the multilevel

theory of hierarchical decision making and identified the internal and external factors influencing the decision accuracy of the team. Three components; informity, validity and sensitivity are the core constructs that carry the relationship between environmental variables and accuracy of decisions (Hedlund, Ilgen Hollenbeck, 1998). Initially, decision informity is the total amount of available information on cues.

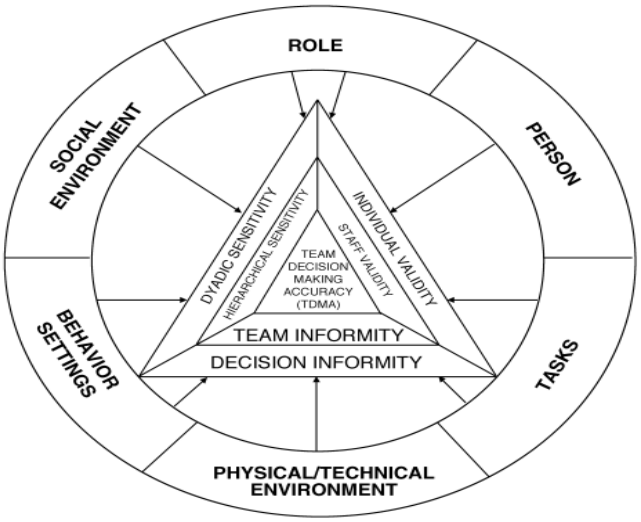


**Figure 3 - Multilevel Theory of Hierarchical Decision Making (Adapted from Multilevel theory of team decision making: Decision performance in teams incorporating distributed expertise," by J. R. Hollenbeck, D. R. Ilgen, D. J. Sego, J. Hedlund, D. A. Major, J. Phillips, 1995, Journal of Applied Psychology, 80, p. 299).**

Validity is measured by the correlation of individual judgments and the leader's decision. Whether a member's judgments are parallel with the leader is important for assessing the contribution of individuals. Hierarchical sensitivity is leader's consistency on preference of weight assigned to each member's judgment. Correlation between the recommendations plays an important role on hierarchical sensitivity. On their review, Humphrey and his colleagues conclude that advice utilization literature analyzes the factors that affect hierarchical sensitivity. The factors

shown in figure 3 are investigated through judge-advisor systems. In other words, advice taking framework contributes to understand the dynamics of team decision making. By this way, scholars aim to explain the variance in decision making and decision accuracy.

Since advice taking literature benefits from the findings of group decision making literature, studies on group performance may give a hint to the necessity of aggregating opinions, because similar variables take role on the experimental design of two frameworks. There has been considerable amount of research on comparison of individual and group performance since the beginning of 20th century (e.g., Ray, 1955; Zander, 1979). Scholars agreed on the conclusion that individual judgments are less accurate than the group's in average (Einhorn, Hogarth & Klempner 1977; Surowiecki, 2005). Most studies measured the effects of various factors on the performance of individuals and groups in a problem-solving framework. These factors can be gathered around four headings; task, process, individual differences and methodology (Hill, 1982). Almost similar factors are examined through literature review of advice taking and utilization (Bonaccio & Dalal, 2006).

## 2.2 What is advice and advice utilization?

The formal definition of "Advice" is "words given or offered as an opinion or recommendation about future action or behavior" (The Oxford Dictionary and Thesaurus, 1997, p. 24). In organizational context, people take advice in case of solving a problem or deciding on some issues related to the work. In JAS framework, advice is given in form of estimation or choice on a problem. Nevertheless, Bonaccio and Dalal (2006) criticized the scholars as

“research on advice giving and taking has itself paid insufficient attention to defining the term: advice” (p. 143). In her book, Gibbons expands the definition of advice (2003). She illustrated the definition of advice as the acquisition of emotional support and assistance, recommendation of new reasoning and alternative or confirmation of making the initial choice.

Another important term in this framework is the advice utilization. It is defined as the degree to which the decision maker’s opinion is changed by the advice. The opposite of advice utilization in the literature is advice discounting which refers to the underweighting of advice in a decision. It is trivial that unutilized advice does not make any sense and so loses its meaning. Thus, advice utilization is an important concept that scholars examine the factors affecting it (e.g., Sniezek, Schrah and Dalal, 2004). Besides that, most studies support that people discount advice due to different reasons. Whether an advice is utilized or discounted can be measured by quantitative analysis. There are different methods of measurement for advice utilization or discounting that are presented later, but in general it is the average absolute difference between initial opinion of decision maker and advisor’s decision (Van Swol, 2009).

## 2.3 Types of Advice

Decision making researchers typically study advice taking design in which an advisor gives specific recommendation about the decision maker's problem in form of what should be done. However, there are also different types of assistance in decisions (Dalal & Bonaccio, 2010). The aforementioned framework for advice is called "Recommend For" whereas the authors introduce the opposite recommendation type referred as

"Recommend Against". As an illustration, advisor may recommend not to choose particular alternative/s which leads the decision maker to elimination of some paths. These two types of advice are based on advisor's personal judgment, so accuracy or consistency of the advice is not proven.

Rather than judgmental advice, "information", the advisor has, may assist in reaching to the decision. This may be done through giving final decision or eliminating an alternative. Another way of assistance is to show how to decide on existing opinion and recommendations. In other words, advisor can provide the degree of advice utilization such as choosing the average of the recommendations or giving more weight to the initial opinion. Beside all these types of assistance, advisors may provide socio-emotional or social support that influences the decision maker. The reason why people take such assistance is that in important decisions people mostly could not manage with the emotional distress. It is revealed that women give more reaction to the social support which confirms Basow and Rubenfeld's (2003) finding that women weight advice in form of social support more than man.

Dalal and Bonaccio (2010) compared these types based on the advice utilization and results indicate that decision makers mostly prefer to get information and then advisor's preference on the problem (Recommend For). It is also concluded that providing combinations of different types of advice contributes decision accuracy of the judges. Although the portion of different types of advice (apart from "recommend for") cannot be ignored in real life setting, almost all JAS studies follow the typical advice type in which advisor provides his/her opinion about a specific problem. Therefore hereafter, "recommend for" is used for the term "advice" in explaining different approaches and analysis.

## 2.4 Motives behind Advice Taking

Why do people feel the need to seek for additional opinions before deciding on a problem or issue? In psychology, the seminal work of Brehmer and Hogafors (1986) mentioned the need for expert advice in complex decision making situations. First of all, real decision problems are different from textbook problems in a sense that there are no strict and clear paths, cues or options. This increases the complexity and human judgment becomes insufficient to find the optimal solution. Another reason is that people intend to enhance their perspectives (Yates et al., 1996), look to the problem in new ways (Schotter, 2003) and so increase their accuracy on a problem (Yaniv, 2004a; 2004b), so they involve in interactions with others, especially ones that have expert opinion.

Researchers touch to this issue from cognitive and social perspectives as well. "Two heads are better than one" is the appropriate aphorism that shows the motive behind seeking advice. Kennedy and his fellows conducted research on accountants and it is shown that advice is solicited to increase self-respect and support their reasoning for the decisions (Kennedy, Kleinmuntz, and Peecher, 1997). Furthermore, Harvey and Fischer (1997) analyzed the reasons for taking advice by providing training in case of judgments that differs in importance. They found that people do not prefer to avoid totally rejecting advice even if it belongs to less trained judges. This is most probably because combining decisions result in reduced errors. Moreover, social compliance may be another reason for the reluctance to avoid advice (Sniezek and Buckley, 1995).

The second reason for soliciting advice is to improve the judgment quality by the opinions of more trained advisors. This is a trivial result, but studies show that even free and correct advices are not fully utilized by the decision makers (e.g., Gardner and Berry, 1995). The last reason for advice utilization is revealed from an experiment in which the importance of judgment differs. Results indicate that people are more conscious on high-risk judgments and so they intend to take advice twice more than the low-risk judgments case. Harvey and Fischer (1997) concluded that people share their responsibility with others by taking advice in risky decisions.

## 2.5 Judge-Advisor System (JAS)

Sniezek and colleagues introduced the Judge-Advisor System (JAS) paradigm which is about how judges weight the recommendations of advisors (Sniezek & Henry, 1989; Sniezek & Henry, 1990). In JAS paradigm, one or more advisors make judgments on a problem and give their recommendations on the decision. Then judge faces the advice of advisors and makes the final decision based on his initial opinion and advice. While giving the final decision, judge considers the expertise of advisors and there is a social influence on the judge as well (Messick & Ohme, 1988). Judges mostly believe that advice helps to give more accurate decisions and so they may overweight the advice given. Especially in "choice" tasks, there exist two or more advices that are completely different from each other and the judge do not have an option to combine those recommendations, so chooses the one. A typical example for this is the choice of new product development (e.g., Schrah, 2000; Schrah, Dalal & Sniezek, 2006).

Sniezek and Buckley (1995), the founders of this model give the review of a paper as an example for the judge-advisor framework. In reviewing an article, the editor becomes the judge and the reviewers are the advisors. The editor is responsible for the decision, but consults to the reviewers' opinion. Then the editor weights each advice and gives the final decision. In general, Yaniv (1997) asserted that when a judge is presented with recommendations from multiple advisors, s/he gives weights to each advice for the final decision. In giving these weights, judge may assign positive weights to all advices (weighting) or a recommendation may be weighted zero (trimming) in order to reduce the number of advices. The next section concerns with the factors that influence the utilization of advice.

## 2.6 Factors affecting Advice Utilization

As it is illustrated, JAS system dominates the most real-life situations and so it takes the attention of scholars for two decades. Especially the introduction of the JAS concept directed the scholars to concentrate on the dynamics of judgment in an advice giving/taking environment. On one hand, individual level variables such as judge-based factors take role on the advice utilization. Specifically, the effect of judge's initial opinion and attributes about the problem are examined. Moreover, judge's confidence is analyzed by considering its relation with credibility and consistency of the advisor. On the other hand, experimental factors such as price or reward options, requisition of advice and complexity of tasks are mentioned. The next section provides an analysis of dynamics between the judge and advisor.

### 2.6.1 Judge and Advisor based factors

#### 2.6.1.1 Credibility of the source of advice (expertise of advisors)

In social psychology, most studies discuss the credibility of the expertise in social interaction and expert power is one of the six bases of power that occur during social influence (French & Raven, 1959; Raven, 1965). Expert power results from the faith that the expert has better insight or knowledge than the judge about what decision is best under the circumstances. This is one of the reasons why people solicit advice in most decisions as mentioned before. There are different perceptions of expertise in JAS literature such as demographic features and experience (Feng & MacGeorge, 2006) or problem-related knowledge (Goldsmith & Fitch, 1997). The level of expertise influences the source credibility in a positive or negative way (Birnbaum & Stegner, 1979). Birnbaum and Stegner (1979) carry out an experiment on the different sources of advice by asking the price of used cars. They assign sources as with low, medium and high expertise. It is found that judges rely on high expertise and so weight it more than other sources.

Reliability of the advice provider takes a big part on the decision process. However, some contradictions may occur during advice taking even if the source is reliable. For instance, Goldsmith and Fitch (1997) made an interview with young people about advice taking sessions. They found that young people see their parents' recommendation as compulsory advice resulting in negative reactions. They concluded that people perceive advice from an expert as less compulsory and more beneficial. On the other hand, some studies compare the effect of expert advice and novice advice on decisions (e.g Jungermann and Fischer, 2005; Meshi, Biele, Korn and Heekeren, 2012). Especially Meshi and his friends concentrate on the brain

activity during utilization of expert advice. By controlling the signals in particular parts of brain, they found that judges value expert advice more than novice advice. Moreover, Sniezek and his colleagues (2004) measured the advice accuracy by varying expertise of the advice provider and monetary reward allocation. They noted that expert advice is more accurate than novice advice and this increases the accuracy of the judges taking and utilizing expert advice.

#### 2.6.1.2 Confidence expressed by parties

Confidence is defined in JAS framework as judge's reliability and trustworthiness on him/herself or the advice givers, as well as advisor's self-assurance. It is measured via asking advisors and judges what percentage you are confident about the decision or by using direct rating method. Both advisor's and judge's confidence has been investigated by the scholars. Initially Sniezek and Henry (1989) argue that confidence and quality of decisions have equal importance in decisions. However, this situation causes "confidence heuristic" that is relied on by the judges (Price and Stone, 2004). Since judges utilize advice with respect to the expertise and knowledge of the advisors, they first consider the confidence of the advisors as if it reflects the characteristics and this heuristic leads judges to give overconfidence on the advisors. Nevertheless, judges substantially rely on the confidence of the advisors and confidence estimates are used to measure advice utilization (e.g., Kuhn & Sniezek, 1996; Van Swol & Sniezek, 2005).

Judges' initial and post-advice confidence as well as the factors influencing the confidence has widely been studied in the JAS literature. Cooper (1991) measured the confidence of the decision makers in an experimental design,

provided a range of advice from multiple advisors and then asked them to use piece of advice as much as they want. Results indicate that pre-advice confidence level is negatively related to advice utilization. In other words, more confident judges solicit less advice. Studies conducted by Budescu and his colleagues primarily focus on the factors influencing the decision maker's confidence (e.g., Budescu & Rantilla, 2000; Budescu, Rantilla, Karelitz & Yu, 2003). They observed that the more relevant cues and shared information, higher confidence decision makers show (see Sniezek & Van Swol, 2001).

#### 2.6.1.3 Accuracy and consistency of advisors

In one judge-one advisor setup, advisor gives an opinion about the problem and then judge finalizes the decision based on the own choice and advice received. However, things get complicated in case of two or more advisors exist due to the discrepancy of the advisor's judgments. It is particularly important to note that accuracy of the advisors is one of the core variables that influence judge's decision accuracy (Hedlund et al., 1998; Humphrey et al., 2002). Accuracy of the advisors depends on the task-relevant information and judges are mostly aware of this fact, so they can eliminate bad advisors. Nevertheless in most cases, judge is not allowed to interact with the advisors and this leads to ambiguity about weighting the advisors who have distinct opinions. In fact, agreement of advisors is directly related to judge's confidence in a sense that judge gives decisions more confidently unless advisors disagree with each other (Savadori, Van Swol, & Sniezek, 2001).

Research has shown that judge's confidence is low in existence of disagreement between advisors accessing the same sources of information (Budescu & Rantilla, 2000). In fact, an advisor conflict results in over-

discounting of the advice by the judges (Sniezek & Buckley, 1995). Moreover in case of an extreme advisor, judge's confidence is manipulated in a way that they give more weight to the extreme advice (Budescu & Yu, 2007). This is because of the belief that conflict affects revealed beliefs in positive or negative way during ambiguity (Baillon, Cabantous & Wakker, 2012). Another issue is advisor's information sharing with the judge during decision process. Judges give more weight on the advisors with more unshared information as if they hide decision-relevant information (Van Swol & Ludutsky, 2007).

#### 2.6.1.4 Judge's and Advisor's Orientation

In a JAS setup, judge and advisor/s have different responsibilities and roles. Similarly, they may be faced with distinct motivational factors. It is worthwhile mentioning that the degree of advice utilization highly depends on these factors. In particular, existing research has examined the psychological aspects regarding to judge and advisor/s. From judge's point of view, advice is discounted due to several reasons. First of all, judges are not well-informed about justifications, judgments and reasoning of an advisor resulting in believing in their internal justifications and preference (Yaniv 2004a, 2004b; Yaniv & Kleinberger, 2000). As Yaniv 2004b mentioned; "Individuals are privy to their own thoughts, but not to the thoughts underlying the advisor's opinion…" (p. 2-3). Another reason for advice discounting is the "egocentric bias" presented by Yaniv & Kleinberger (2000). This egocentric behavior is the result of tendency of judges to underweight advisor's opinion relative to their own. Krueger (2003) pointed out that egocentric bias arises even in case of unfamiliar or unstated situations.

Advisor on the other hand plays an important role in decision making as well. However, there have been few studies focusing on the JAS structure in the eye of the advisors (Jonas & Frey, 2003). According to Kray (2000), judges behave according to their own judgments or choice, but advisors give suggestions by considering judge's preference instead of solely holding on that of their own. In other words, people tend to make different choices in advisor and judge roles. Since advisors are not primarily responsible for the outcomes of a decision, it is necessary to question whether advisors face with motivational deficiencies. However, research in this field indicates that it is not the case; contrarily advisors put more emphasis on their decision performance (Kray, 2000). Another reason for this difference is the confirmation bias that judges face in advice taking causing biased interpretation of an advice (Jonas & Frey, 2003).

### 2.6.2 Advice-based factors

#### 2.6.2.1 Single or multiple advice

As mentioned before, typical JAS studies are performed in one advisor-one judge setup (e.g., Harvey & Fischer, 1997). However, there are different ways of getting additional assistance in real life and people do not tend to refuse an additional opinion on a decision. Supporting this fact, most researchers included multiple advices ranging from two to ten advisors (Bonaccio & Dalal, 2006). From the point of researchers, using multiple advisors is beneficial for investigation of quality of advice and its utilization. It is worthwhile mentioning that most studies using multiple advisors attempted to focus directly on different aspects rather than the variation in number of advisors except few studies (e.g., Budescu & Rantilla, 2000; Cooper, 1991).

One seminal finding on JAS literature is that judge's confidence is positively related to number of advice received (Ashton, 1986; Cooper, 1991). Particularly, an increase in the number of advisors leads to a rise in the confidence of decision makers (Ashton, 1986). Budescu and Rantilla (2000) linked this argument to consensus among advisors and found that the number of advisors is effective in case of high consensus within advisors. Supporting this, a different approach has been taken by Cooper (1991) found that less confident judges solicit opinions of more advisors. These arguments indicate a direct relationship between judge's confidence and utilization.

Another study by Yaniv and Milyavsky (2007) noted that there is a proportion that judges give weight to their own opinion (approximately 40%) and this does not change in case of varying number of advisors. This is due to egocentric behavior presented by Yaniv and Kleinberger (2000). Finally, research on the accuracy of recommendations reported that as number of advisors increases, accuracy is also improved at a diminishing rate (Johnson, Budescu, & Wallsten, 2001). On the contrary, one adverse effect of multiple advisors can be information overload which increases the complexity of decisions (Hiftz & Turoff, 1985).

#### 2.6.2.2 Free or Purchased advice

Taking expert advice for free is almost rare in real world. For instance, people pay huge amounts of money to therapists or investment consultants. Supporting this, previous laboratory studies have included financial rewards depending on the decision performance in their experimental designs (e.g., Budescu & Rantilla, 2000; Sniezek et al., 2004). Theoretical approach in this issue stems from the notion that incentives lead to more

effort, but not necessarily better performance (Camerer & Hogarth, 1999). Furthermore, the tendency to avoid from being wasteful is another factor for the underlying idea of including financial rewards (Arkes & Blumer, 1985). Some studies in the JAS literature have used incentive methods for motivational factors (e.g., Yaniv, 2004b) whereas some tested the effect of reward on advice use and accuracy (e.g., Dalal, 2001).

It is worthwhile mentioning that the cost of advice plays an important role on advice utilization especially when the judge has reward power (Sniezek & Van Swol, 2001). Gino (2005; 2008) linked this to the sunk costs fallacy presented by Arkes and Blumer (1985) in a sense that paying for advice leads judges to tend to use advice in order not to be wasteful. Another perspective on this issue is that credibility of advisor increases as it is paid for expert advice (Patt, Bowles & Cash, 2006). Research has also focused on the effect of financial rewards on advice discounting rather than utilization. Particularly, two studies propose opposite arguments on this issue. Sniezek et al. (2004) revealed that use of advice and decision performance increases in case of financial reward. In contrary, Dalal (2001) established prisoner's dilemma situation and found that judges do not intend to cooperate with the advisors and so, discount advice more in case of financial incentives.

#### 2.6.2.3 Requested or Unrequested advice

Another psychological perspective of advice taking is whether imposed advice causes threat-to-self-esteem is introduced by Fisher and his colleagues (Fisher, Nadler & Whitcher-Alagna, 1982; Nadler & Fisher, 1986). The threat-to-self-esteem model indicates that receiving advice that threatens the freedom of choice of the judge will lead to negative outcomes

such as over-discounting the advice. The main reason for such an outcome is that receiving advice may imply that the judge is inferior or incapable to decide on the task. Negative reactions of the decision maker can also be illustrated by worse decision performance, reduction in motivation and effort, as well as decrease in advisor's confidence ratings (Newsom, 1999).

In the line with above reasoning suggested by psychologists, JAS scholars pay attention to the imposition of advice. Those studies enabled judges to request to receive recommendation of specific advisors (e.g., Gardner & Berry, 1995). Studies agreed upon the result that unrequested advice is regarded as intrusive (Goldsmith, 2000) and inconvenient (Deelstra et al., 2003). Gibbons, Sniezek and Dalal (2003) found that decision makers do not always request advice and judges discount the unrequested advice more than requested advice. This is true even if judge is aware of the accuracy of an advisor (Gardner & Berry, 1995). In short, advice utilization is highly dependent on advice requisition and unrequested advice may lead to threat-to-self-esteem and so over-discounting.

#### 2.6.2.4 Real or Hypothetical Advice

Most JAS scholars use real advice in experimental design of their studies (e.g., Van Swol & Ludutsky, 2007; Dalal, 2001). However, few studies use real advisors or experts due to the unavailability and time limitation. Instead, experimental subjects are assigned as judge or advisor. Then advisors are enrolled in sessions to take task-relevant trainings. This kind of design helps researchers to create different variations and explore its potential influence (e.g., Gardner & Berry, 1995). For example, most studies have investigated the interaction between judge and advisor (will be

mentioned later). Furthermore, advisor's recommendation and judge's decision are simultaneously performed due to measure the confidence of both parties as explained before.

On one hand, some studies use advice that is gathered before being presented to judge (e.g., Schrah et al., 2006; Yaniv & Kleinberger, 2000). The advice may or may not be from a real advisor, but it is not announced to the decision makers since it has no relevance for the purpose of a study. On the other hand, advice can be hypothetically formed via software by the experimenter. This situation is hidden to the judges in some studies, in other words they are believed that there exists real advisor/s (e.g., Brehmer & Hagafors, 1986). Furthermore, some studies use this type of advice although judges are not informed about the source of advice whether it is real or not (e.g., Fischer & Harvey, 1999; Harvey et al., 2000) whereas some studies told the truth (e.g., Budescu, 2006). According to Budescu (2006), informing judges about the hypothetical advice may result in sacrificing realism about the credibility, but it establishes high level control over advice.

#### 2.6.2.5 Random or Perfect Advice

In a JAS experiment, computerized advice is used for the several reasons. Initially, it is practical and time saving to prepare estimations in numerical or multiple choice tasks. Furthermore, randomization makes it enable to convert computerized advice into perfect, misleading, expert or non-expert advice. Perfect advice consists of true estimates whereas misleading advice is introduced after judge's initial estimate in the opposite side (true estimate falls between advice and initial opinion). Another illustration for the practicality of computerized advice is that the experimenter (e.g., Gardner &

Berry, 2006) can measure the effect of optional and compulsory advice in an easy way by manipulating the computer software.

Randomized advice can be in form of expert or non-expert advice depending on the random error allowed. In other words, expert advice can be generated by setting a little random error and variance whereas the opposite condition holds for the non-expert advice. In fact, some studies compared the influence of real and statistical advice and found that people pay greater attention to the real advice than statistical advice (Önkal, Goodwin, Thomson, Gönül & Pollock, 2009). Nevertheless, randomized advice leads scholars to investigate inner mechanisms in a quick way.

There are different practical uses of random advice in the literature. Harvey and his colleagues (2000) added random noise in different amounts to the actual estimates and so they created hypothetical advisors that differentiate in quality. Furthermore, Harvey and Fischer (1997) used random advice to categorize two types of advice as correct advice and incorrect advice. Some studies randomly arranged the advice with respect to the position relative to the pre-advice estimate (e.g., Yaniv, 2004b). Björnsson, Hafsteinsson, Johannsson and Jonsson (2004) compared the learning progress of people playing a game provided by helpful advice and random advice. Results show that both helpful and random advice lead to similar performance.

### 2.6.3 Other Process-based factors

#### 2.6.3.1 Existence of pre-advice opinions

Judges are usually asked to give pre-advice estimates about decision task before they are informed about the advice. However, most studies do not

allow DM to announce their initial opinion about the decision task (e.g., Harvey et. al., 2000; Sniezek & Van Swol, 2001). The reason is that the purpose of the study is not related to the advice cues but advice quality or confidence issues. On the other hand, Gibbons and his colleagues (2003) used initial decisions by presenting them to advisors. Results showed that advisors are more eager to offer their opinion in case of conflict with the decision makers. Aside from that study, most scholars used pre-advice decisions to measure the utilization and decision performance of the judges. In fact, the main motivation behind JAS studies is how people use advice to reach more accurate decisions.

Sniezek and Buckley (1995) support that pre-advice decisions influence the degree of advice utilization. In line with this reasoning, studies considered the relationship between initial decision and advice. Consistency between these two decisions is seen as a significant factor in advice taking (Yaniv, Choshen-Hillel, & Milyavsky, 2009). It is found that more consistent advice is preferred more since it confirms judge's opinions (Savadori, Van Swol, & Sniezek, 2003). Supporting this, judges tend to give more weight on advice that is consistent with initial opinion (Yaniv, 2004a, 2004b). Not only consistency, but also similarity of decision process leads judges to be influenced more from advisors that have similar decision process (Schotter, 2003). In short, judges tend to trust the advisors that act in a similar manner during decision process.

#### 2.6.3.2 Interaction between judge and advisors

For convenience and time concerns, most JAS experiments do not allow interaction of the members. The collection of the decisions and advice are

also not simultaneously performed. Advice is collected before judge's decision period or generated through computer (e.g., Schrah et al., 2006). Judge and advisors interact with each other through written documents in different places (e.g., Sniezek et al., 2004), again written form in the same experimental room (e.g., Dalal, 2001), web-conferencing (e.g., Gibbons et al., 2003) and face-to-face (FtF) communication (Savadori et al., 2001). In all cases, judge interacts with the advisor and is informed about the decision process of the advisor.

A great deal of research has been devoted to the comparison between FtF and web-conferencing systems especially in group decision literature. Hedlund and his colleagues (1998) found that FtF interaction increases judge's decision accuracy more than computer-aided interaction. On the contrary, Handgraaf, Milch, Appelt, Schuette, Yoskowitz and Weber (2012) investigated the performance of decisions through face-to-face and web-conferencing communication. They show that there is no difference between two systems in terms of performance and web-conferencing leads to more efficient use of time.

#### 2.6.3.3 Type and Complexity of tasks

JAS experiments typically involve decision tasks in the form of questionnaires to be answered by the decision makers by utilizing additional opinions. Most studies identify their decision task in accordance with the subject's background and characteristics as well as the purpose of the study. By considering the outcome derived from the experimental designs, Billings and Scherer (1988) divided decision tasks into two types such as choice tasks and judgment tasks. Choice tasks consist of multiple choice questions that

have alternatives for the answer of each question in the survey. After receiving advice, decision maker is responsible for choosing an alternative for each question. Studies consider whether the initial choice is changed or not (e.g., Sniezek & Buckley, 1995).

Choice tasks are preferred less than judgment tasks. The first reason is that judgment tasks involve quantitative decisions that decision makers perform by estimating probability, score or date of certain events. Numeric estimations can be easily changed by offering advice whereas it is more difficult to change people's preferences in multiple choice tasks. Kruger, Wirtz and Miller (2005) support this view and conclude that people generally resist changing their preferences on choice tasks even if they know that it will increase their performance. This may be due to the egocentric act that changing choice is seen as lack of confidence, but this is not the case for judgment tasks since changing estimation is seen as an ordinary action.

Furthermore comparing to the choice tasks, judgment tasks produce healthier results to measure advice utilization since how a decision maker has deviated from initial decision can be calculated for each decision in judgment tasks. Another advantage of judgment tasks is that they use interval estimates in addition to point estimates. By this way, experimenters can measure the upper or lower boundaries and probability of capturing true estimations. There are lots of studies using judgment tasks in their experiments. For instance, such studies used price (Sniezek et al., 2004) or sales volume (Harvey et al., 2000) of certain products, probability of occurrence of an event (Budescu & Rantilla, 2000), outcome of a natural disaster (Harvey & Fischer, 1997), date of historical events (Yaniv & Kleinberger, 2000) or stock price forecasting (Onkal et. al., 2009).

Complexity of decision tasks plays an important role for the purpose of a JAS study. A complex decision task is observed when judge is unfamiliar with the task, there exist lots of alternatives and paths, or too much information is necessary to decide. Complexity of a task for DM entails additional effort for decision making. In such conditions, DM may dislike the decision task and lose motivation which results in an unhealthy and biased data collection. Besides, complexity of the task may hinder the observation of the desired results. For instance, an experimenter considers the effect of interaction between JAS members, but the decision task is so complex that advice is utilized at extraordinary levels for each situation. Payne et al., (1993) support this argument that DM's change their decision strategies during complex tasks. In short, choosing an appropriate decision task for a JAS study is essential and complexity of a task should be adjusted according to the purpose of the study.

## 2.7 Techniques on measurement of Advice Utilization

The aforementioned studies explored different dynamics of advice taking and aimed to see which factors play role on judge's preferences and decision accuracy. For instance, scholars take into account many aspects of an experimental design such as the presence of pre-advice decision, number and qualification of the advisors and the interaction between advisor and the decision maker. In what degree these factors influence the advice taking process of a judge is examined by using different techniques. The goal is to find the degree of advice utilization of a decision maker which is a measurable concept used to operationalize 'advice taking'.

Different measures of advice utilization have been developed since 90's (Bonaccio & Dalal, 2006). The data produced by the two types of decision task (choice and judgment) are measured by different techniques, because the outcomes derived from a choice task is true or false options whereas judgment tasks involve some numeric values. Gibbons and his colleagues (2003) used choice tasks and called the decision that judge follow advice rather than own choice as "shift". After excluding the case in which all initial choice and advice are same, "shifts" shows the degree of advice utilization.

On the other hand, three main approaches exist in measuring the advice utilization on judgment tasks. The first approach uses various formulas to measure the judge's utilization of advice (Harvey & Fischer, 1997; Yaniv, 2004; Yaniv & Kleinberger, 2000). Second, advice utilization is measured by regressing the judge's final decision on advice and/or the initial decision (Harvey et al., 2000; Hedlund et al., 1998; Hollenbeck et al., 1995). Lastly, direct rating approach is used through implementing Likert-type scale. All three techniques will be explained in a detailed way in next section.

### 2.7.1 Formula-based approaches

JAS studies usually involve measurement of the degree of advice utilization in existence of certain variable such as financial reward. To do so, they form two subject groups in which one gets financial assistance depending on the decision performance whereas the other group is not paid for success or failure. The first approach to investigate the degree of the advice utilization in such a case is making use of formulae. All subjects of the experiment are asked to give their initial and final estimate after receiving advice, so elements of these formulae are initial and final estimate and the advice.

Out of the many suggested ones, three of them gained popularity through different studies. The first one is devised by Harvey and Fischer (1997). The researchers named their formula as "Advice taking" represented in Equation 1. The numerator is the difference between judge's final estimate and initial estimate indicating the amount of change in judge's decision. The denominator is the difference between advice and judge's initial estimate which is indicator of judge's initial accuracy. Thus, the ratio represents to what degree initial estimate is changed towards advice. As an illustration, assume that the advice is 100; initial estimate is 60 and final estimate is 90, then "advice taking" becomes 0.75 meaning that 75% of the advice and 25% of the initial decision is combined for the final decision. One should note that "advice taking" may be negative or more than 1 depending on the orderings of estimates and advice.

**Eq. 1:** $$\textbf{Advice taking} = \frac{\textbf{judge final estimate} - \textbf{judge initial estimate}}{\textbf{advisor recommendation} - \textbf{judge initial estimate}}$$

Yaniv (2004b) took the absolute value of both the numerator and denominator of the Harvey and Fischer (1997)'s formula and presented the weight of advice (WOA) seen in Equation 2. Taking the absolute values is due to eliminate negativities in a data, but in order to get negative result in "advice taking" formula, judges should move away from advice on behalf of their initial estimate. Existence of such situations in a data hurt the results since an unutilized advice become over-utilized after taking absolute value of the ratio.

**Eq. 2:** $$\textbf{WOA} = \frac{|\textbf{judge final estimate} - \textbf{judge initial estimate}|}{|\textbf{advisor recommendation} - \textbf{judge initial estimate}|}$$

The third advice-taking measure is developed by Yaniv and Kleinberger (2000). They changed the numerator as the absolute difference between advice and the final decision and named it as weight of own estimate (WOE)

as seen in Equation 3. According to the previous example, WOE becomes 0.25 (WOA was 0.75). 0.25 is judge's utilization rate of initial estimate and so it is the degree of advice discounting. This leads to the result that WOA and WOE add up to 1 and give results in terms of opposite measures.

**Eq. 3:** $$\mathbf{WOE} = \frac{|\mathbf{advisor\ recommendation - judge\ final\ estimate}|}{|\mathbf{advisor\ recommendation - judge\ initial\ estimate}|}$$

These formulae similarly indicate the proportion of adjustment in decisions with respect to the distance from the advice. It is evident that, if the final estimate is equal to the advice, then the result is undefined since the denominator becomes zero. Final estimate being equal to initial estimate means that advice taking ratio and WOA take the value of 'zero' but WOE takes the value of 'one' which indicates that no advice is utilized. If the advice is equal to the final estimate which is not equal to the initial estimate, then WOE takes the value of 'zero' which shows perfect utilization. In short, these three measures give similar results from different directions, but "advice taking" gives the true values in bracketing case and so it is preferred in analysis of advice utilization.

### 2.7.2 Regression-based approaches

Another way of advice utilization measurement is regressing the judge's final estimate on the advice (e.g., Harvey et al., 2000; Hedlund et al., 1998) and pre-advice estimate if exists (e.g., Lim & O'Connor, 1995). Most studies that used regression-based approach focused on understanding how individuals incorporated the advice from multiple sources in their final decisions. It is possible by including the different advisors' recommendation in the model. The regression model is expressed in Equation 4. The coefficients represent the utilization from initial estimate and the advice/s.

**Eq. 4 :Final Estimate = $\beta_0$ + $\beta_1$*Initial Estimate + $\beta_2$*Advice$_1$ +…+ $\varepsilon$**

One assumption behind this approach is that judge's initial estimate and advisors' recommendations are exogenous whereas judge's final estimate is endogenous (Bonaccio & Dalal, 2006). In their review, Bonaccio and Dalal mentioned two utilization indices that measure the effect of advice as a predictor. The "usefulness index" (Darlington, 1968) is the percentage increase in criterion variance explained when an advice as a new predictor is added to the model. On the other hand, dominance weight is proposed by Azen and Budescu (2003) and is used to determine the importance of variables by aggregating the variance contribution of all pairs of variables including all the subsets of the predictors. It is mainly the average percentage increase in criterion variance explained when all subsets of predictors are included in different models. It is seen that the dominance weight is the association of each usefulness index calculated in each subset.

### 2.7.3 Direct rating approach

Direct rating is a widely used method for measuring the level of different characteristics or capabilities. This method uses likert-type scale and requires the decision makers to explicitly rate the extent they believe to have utilized the advice. A typical example for direct rating approach is that a customer tastes a new product of your company and you get feedback from that customer by asking whether s/he likes the product. The customer is asked to choose a number from the scale such as 1-10 in which 10 means that the customer loves the product whereas 1 indicates full of hate. It is worthwhile mentioning that each number must have a clear definition that fits the whole scale. Studies mostly use 3, 5, 7 and 10 point scales depending on the response categories needed. However, 5 and 7 point scales give mean

values that are closer to the actual mean of the data and so perform better than other scales (Dawes, 2008).

JAS studies use Likert scales in order to measure confidence and advice utilization of the members. Specifically advisors and judges are asked to choose a point in a scale in order to show the degree of confidence or assistance received during decision. How to measure advice utilization by this approach is progressed as this; in a 7 Likert scale study that contains 20 questions, if it is asked to the subject for each question that "To what degree, you utilized the advice", then the mean of all 20 answers divided by 7 is the degree of advice utilization based on this approach. Direct rating is seen as an efficient way of measuring people's behavior in the JAS literature. Van Swol and Sniezek (2005) found that there is a significant relationship between confidence ratings and the degree of advice utilization. However, it is found that the confidence rating is not always an indicator of advice quality (Philips, 1999).

## 2.8 Techniques on Measurement of Advice Accuracy

In order to calculate whether people consult an additional opinion or not, scholars use various measurement techniques as mentioned in the previous part. The novel reason for measuring advice utilization is to find out the improvement in decisions after receiving an advice. However, it is clear that it does not make any sense to measure the improvement in decision accuracy of a person who prefers to refuse receiving advice). In that rare case, there is no reason for decision makers to change their judgments. Nevertheless, research on advice taking indicated that people often react to an advice in a positive or negative way. The question here is that whether

people can improve their decision performance after reacting to an advice or whether utilized advice leads to better decisions than the discounted advice.

Accuracy is defined as how well a new estimate (assisted by an advisor) is able to get closer to the true estimate (Makridakis & Wheelwright, 1989). All JAS studies, that measure the advice accuracy, include judge's initial estimates. Increase in decision performance can be measured by comparing pre-advice and post-advice estimates with the true estimates of the decision task. In doing so, it is necessary to consult to parallel research area which is forecasting (particularly supply-chain forecasting and demand forecasting). Researchers and practitioners show different approaches in measuring the forecast accuracy of an analyst or an institution. Most studies contradict on the reporting of forecast error vs. forecast accuracy (e.g., Hawitt, 2010; Hoover & Little, 2009). It depends on the purpose of a study, but both measurement styles give similar results. In particular, scholars conduct regression of two data sets in which one of them is true data and the other is estimated or forecasted.

As Harvey and Fischer (1997) asserted; "Selecting measures of forecast accuracy and advice taking is not a simple matter" (p. 121). That is because different types of data require choosing the appropriate accuracy measure to obtain the best results. Makridakis, Wheelwright and McGee (1983) are one of the earliest forecast experts that focus on the accuracy measures. They studied forecasting methods and test the forecast accuracy by statistical evaluation measures that use forecast error. Studies recommended the use of various measures up to recent times and they are categorized into two main types as scale-based and percentage-based measures. Here, let n be the number of judges as subjects and $e_t$ be the error term that is the difference

between true estimate ($Y_t$) and judge's estimate ($F_t$). Percentage-based measures are based on percentage error ($p_t$) that is the ratio of error term, $e_t$ and true estimate ($Y_t$) multiplied by 100. Below are the most commonly used accuracy measures and their formulae;

**Scale-Based Errors**

MSE = Mean Square Error = $\mu(e_t^2)$

RMSE = Root Mean Square Error = $\sqrt{MSE}$

MAE = Mean Absolute Error = $\mu|e_t|$

MdAE = Median Absolute Error = $\mathrm{md}|e_t|$

**Percentage-Based Errors**

MAPE = Mean Absolute Percentage Error = $\mu|p_t|$

MdAPE = Median Absolute Percentage Error = $\mathrm{md}|p_t|$

RMSPE = Root Mean Square Percentage Error = $\sqrt{\mu(p_t^2)}$

RMdSPE = Root Median Square Percentage Error = $\sqrt{\mathrm{md}(p_t^2)}$

One of the seminal studies on this issue found that RMSE is the most widely used accuracy measure by forecast experts (Carbone & Armstrong, 1992) whereas MAPE is found to be most reliable and valid measure (Armstrong & Collopy, 1992). Furthermore, there are also relative error measures such as MRAE (mean or median relative absolute error) (Armstrong & Collopy, 1992) or more recently MASE (mean absolute scaled error) (Hyndman & Koehler, 2006) that are used by forecasters in order to eliminate the scale difference between data sets. Scaled error in MASE measure is calculated by naïve forecast method that is valid in time-series data.

Due to this scale problem, some studies criticize MAPE (e.g., Makridakis, Wheelwright & Hyndman, 1998; Swanson, Tayman, & Bryan, 2011). Recent studies introduced new measure in order to be alternative for MAPE. One of them is MAD-to-Mean (or MAE-to-Mean) founded by Hoover (2006) that is calculated by MAE divided by the mean of data. It gives the average error as a percent of the average volume, so scale is not reflected in the data. Kolassa and Schütz (2007) called MAE-to-Mean as a weighted average of absolute percentage errors. Finally, Hoover (2009) recommends that MASE and MAE-to-Mean are the best alternatives for MAPE.

Note that forecasting and advice taking literatures have some similar and different characteristics in terms of data sets used in experimental designs. The difference is that experimental design of JAS studies involves cross-sectional data that does not vary with respect to time whereas forecasting data is usually a time series data. Thus, not all forecast accuracy measures are relevant for the JAS studies especially that are used for time-series data (Harvey & Fischer, 1997). JAS scholars mostly use MAE (e.g., Soll & Larrick, 2009; Yaniv & Milyavsky, 2007; Harries et al., 2004) and MAPE (e.g., Yaniv, 1997; Harvey & Fischer, 1997). The similarity of two literatures is that JAS experiments also produce data sets that can have different scales. Therefore, this point is important in choosing an appropriate accuracy measure in analysis.

# CHAPTER 3

# METHOD

## 3.1 Categories of Adjustment

Harvey and Fischer (1997) asserted that the ratios for measuring advice utilization mostly fall between zero and one. In their review, Bonaccio and Dalal (2006) mention this fact and define the values that are not between zero and one as "out of range" values that likely designate problematic cases. However, a problematic case in this sense might also represent a situation where a judge's final estimate is oppositely affected by the advice and may, in fact, illuminate another facet of the advice taking phenomena. Thus they strongly suggest scholars to focus on these problematic cases (Bonaccio and Dalal, 2006). In this study, I try to answer this call by categorizing our data into three parts with respect to the position of the advice, the initial estimate and the final estimate. These parts can be described as extreme adjustment, bracketing and moderate adjustment.

In extreme adjustment, advice falls between the judge's initial estimate and the final estimate. This occurs when the judge totally changes his/her idea in a positive way after the advice is shown. That is when judge's final estimate overshoots. In bracketing, judge's initial estimate falls between the advice and the judge's final estimate which happens when the judge changes his/her idea in the opposite direction of the advice. In this situation, judge's final estimate moves away from the advice. Lastly, moderate adjustment

represents the typical situation in which final adjustment falls between the advice and the initial estimate. This case can be anticipated since judges mostly give positive weights to their own opinion and the advice.

## 3.2 Participants

In a "one judge, one advisor" setup which is the typical framework of JAS studies, I prepared an online questionnaire in which the subject as a judge is first required to provide a pre-advice estimate on a general knowledge item. And then he is presented with an advice. After examining this advice, he is asked to give a final estimate for the item by reminding the initial estimate and also a rating of the degree of advice utilization. Since there is no time and space requirement for online surveys, participants were not gathered in a class environment. The participants were not aware of one another and all surveys are completed in two-day period. Participants are voluntarily involved in the experiment and they have not received any financial aid.

A total 66 students participated in the study. The participants were graduate level students from various universities (primarily from Bilkent University and Middle East Technical University (METU)). Thus, they were in age of an interval between 25 and 30. The reason why these participants are chosen is that they have succeed well in national examination and received good higher education and so they are predicted to have higher IQ than an average citizen in our country. This is important for performance on reasoning, judgment and combination of relevant information. Another common characteristic of participants is that technology has been quickly improved in their adolescence era. This leads them to have some idea about knowledge that is shared by public making suitable for the experiment.

### 3.3 Questionnaire and Data

**Table 1 - Sample Question (see Appendix B for survey screen)**

| |
|---|
| Phase 1 (series of 20 questions) |
| In what year was the compact disc invented? |
| Your estimate: ____ |
| Phase 2 (same 20 questions repeated) |
| In what year was the compact disc invented? |
| Your previous estimate was 1990 |
| Advisor's estimate is 1977 |
| Your final estimate: ____ |
| In which scale you utilized advice: ___(1-7) |

Previous research used estimation based on historical events (e.g., Yaniv, 2004b; Yaniv & Kleinberger, 2000), estimating the annual salary of alumni of business schools (e.g., Soll & Larrick, 2004) or estimations of the probabilities of occurrence of certain events (e.g., Budescu & Rantilla, 2000). The current study utilized a similar task and asked the participants to estimate the dates of 20 important inventions that occurred in the 20$^{th}$ century (e.g., internet, ATM machine, heart transplant). It is assumed that people have some idea about the invention period of important materials. Moreover, people can link the invention year of similar products and perform better estimations. The advice is generated randomly with ± 20 years variation from the actual values, because subjects are informed that the advice comes from an expert.

In order to provide a boundary to the estimation task, the participants were instructed that the dates of these popular inventions can be estimated within a range of at most 50 years. The reason why I limited the data for a 100-year

period is that if an invention that is quite distant in the past (e.g., the invention year of steam engine-1698) was used; then the level of variability of the participants' responses would be quite abundant and this would probably cause a bigger mismatch in scale of data sets which is a problem in measuring decision accuracy as mentioned before. It would also be highly probable that the participants may not be knowledgeable about events so far away from the current time and thus may guess wildly or put their faith upon the advice excessively. Supportively, the distance between initial estimate and advice is found to affect the judge's willingness to seek advice (Yaniv, 2004b), so it is logical to restrict the span between pre-advice estimate and the advice.

The decision task consists of 20 questions and each judge's estimates for each invention are considered as an observation that includes initial estimate, advice and final estimate. 66 participants estimated 1320 decision task and in total 1320 observations in the range of 1900 and 1999 as well as 1320 correspondent ratings are obtained. Including the existing data (the advice and true estimates), a data set involves an initial estimate, advice, final estimate, true estimate and a rating score. It is a cross sectional data that includes data sets that have different mean. This is because the invention dates are ranging from 1902 and 1996.

### 3.4 The Setup and Procedure

The reason for preferring online survey is the availability and time concerns of the subjects. As participants do the survey online, there was no interaction between judge and advisor. On the other hand, there was no option to request or refuse the advice. In other words, advice for each decision is

imposed to the judge. Moreover, since no financial incentive is included to the study, advice is freely offered and no reward is designed for the decision performance. Participants received one advice for each decision, so multiple advisors are not allowed for the purpose of the study.

The type of task is chosen in accordance with people's knowledge of general facts so that each subject can have idea about the decision task. In other words, the decision task was not complex and so subjects are not required to deeply analyze the situation. Judges were informed that the advice came from an expert (with good general knowledge in this case) similar to what had been done in existing studies (e.g., Brehmer & Hagafors, 1986; Harvey and Fischer, 1997), but there was no real advisor again similar to previous studies that use hypothetical advice.

It is previously mentioned that accuracy of the advisors influence judge's decision accuracy (Hedlund et al., 1998; Humphrey et al., 2002). Furthermore, Birnbaum and Stegner (1979) found that judges rely on high expertise leading to raise in the source credibility and so weight it more than other sources. Thus, it is aimed in this study that little random error of advice estimates helps decision makers to increase the decision performance. Another issue is that allowing judges to announce the initial opinion is necessary in this study because the degree of advice utilization and accuracy are investigated. To sum up, typical JAS study has been conducted on subjects and initial estimates, advice, final estimates and ratings are gathered to be analyzed in direction of the purpose of the study.

After gathering the data for the three cases, firstly I use formula-based measures and find the utilization degree of judges with respect to Advice

Taking, WOA and WOE. Then I made regression of judge's final estimate on advice and initial estimate and regression outputs are summarized for the three cases. Furthermore, usefulness index and dominance weights are calculated to find the degree of advice utilization. Lastly, mean of ratings for each case is calculated to directly measure in what degree judges utilize advice. The three measurement techniques are compared for the three cases.

After analysis of advice utilization, decision accuracy of judges is measured by using various accuracy measures. Firstly, MAE is chosen as a scale-based measure, because it is widely used by forecast experts and most JAS scholars prefer to do so as well. Secondly, a percentage-based error, MAE-to-Mean ratio is used as an alternative to MAPE. The three cases are compared with respect two MAE and MAE-to-Mean measures within each other and also with the benchmarks. The first benchmark is a constant estimate which is chosen as the mean value of the true estimates. Yates (1994) calls this as constant judge who makes no distinctions among decision tasks.

It is expected from judges to perform better than a constant estimate regardless of the adjustment ordering. The second benchmark is Yates' uniform judge that is the average of initial estimate and the advice. This is previously studied by the experimenters, especially Soll and Larrick (2009) supported that people tend to average their initial opinion and the advice. They found that judges that utilize averaging are more accurate than the others. That is why I prefer to use the averaging method as a benchmark. Finally, decision performances of these two benchmarks are compared with the initial performance and final performance of the three situations

# CHAPTER 4

# RESULTS

## 4.1 Data

After the data is separated into three parts with respect to the position of the advice, I decided to exclude some of the observations due to being equal. Specifically, in 11 observations, advisor recommendation is equal to judge's initial estimate; in 168 observations advisor recommendation is equal to judges' final estimate and in 200 observations judges initial estimates are equal to their final estimates. I left out these cases, since formula based approaches will not produce interpretable results for such cases of equality. In terms of the rest of the data, I have 213 observations for extreme adjustment case; 83 observations for bracketing case; and 626 observations for moderate adjustment case. If only the observations belonging to three cases are considered, 76% of the estimates are in the moderate adjustment condition which is almost contradictory with Bonaccio and Dalal's statement that this condition constitutes; "…overwhelming majority (up to 95%) of observed decisions" (p. 141). Furthermore, judges truly estimate the decision task and do not change their decisions in 17 cases. However, these cases account for very small portion and so it is not worth to consider the effect of true estimations for this study.

## 4.2 Formula-based Measures

In this part, formula-based approaches are compared with respect to three adjustment cases. For all these cases, I calculated the advice taking, WOA

and WOE ratios. Table 2 exhibits the descriptive statistics of the values calculated by the respective formulas; μ: the mean of the values calculated by the formulas, δ: the standard deviation. Bonaccio and Dalal noted that judges usually moves closer to advice and so all these measures fall within zero and one in the moderate adjustment case. Table 2 indicates that on average, judges utilize 69% of advice. Average WOE value also supports the claim in a different perspective that judges discount 30% of the advice. Division of data according to the direction of the adjustment leads to separation of the positive and negative values in advice taking formulae (which is not an issue in WOA and WOE since they are in absolute value form). That is why advice taking and WOA statistics attain the same values in absolute terms.

**Table 2 - Results of Formula-based Measures**

| | **Advice Taking** | | | | **WOA** | | | | **WOE** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | μ | δ | Min | Max | μ | δ | Min | Max | μ | δ | Min | Max |
| Extreme Adjust | 2.5 | 2.5 | 1.0 | 16.3 | 2.5 | 2.5 | 1.0 | 16.3 | 1.5 | 2.5 | 0.0 | 15.3 |
| Bracketing | -1.8 | 2.7 | -17.0 | -0.1 | 1.8 | 2.7 | 0.1 | 17.0 | 2.8 | 2.7 | 1.1 | 18.0 |
| Moderate Adjustment | 0.7 | 0.3 | 0.0 | 1.0 | 0.7 | 0.3 | 0.0 | 1.0 | 0.3 | 0.3 | 0.0 | 1.0 |

As judges extremely adjust their estimates, all observations in three measures become larger than one, because as the judge moves beyond the advice, the difference between final and initial estimate surpasses the difference between advice and the initial estimate. However, in bracketing case, judges move away from advice in different degrees (range is 0.13-17). This results in a higher standard deviation (2.73). Formulation of WOA and WOE leads to the observation that average WOE is exactly one unit less than average WOA in the extreme adjustment case and one unit more in the

bracketing case. WOA and WOE values in the bracketing case are larger than one which shows that judges move away from advice in large amounts on average. An important result is that WOA values indicate high level of advice utilization for the bracketing case, but no advice is utilized in fact. Thus, using WOA and WOE in those cases hurt the reality of results.

### 4.3 Regression-based Measures

In this part, regression-based approaches are compared with respect to three adjustment cases. By regressing judge's final estimate on advice and initial estimate, the degree of advice utilization is measured. The regression equations obtained are:

Extreme Adjustment:

**Final Estimate = 660–0.46*Initial Estimate + 1.12*Advice ($R^2$=0.57)**

Bracketing:

**Final Estimate = 255+ 1.6*Initial Estimate –0.74*Advice ($R^2$=0.87)**

Moderate adjustment:

**Final Estimate = 15+ 0.21*Initial Estimate+ 0.78*Advice ($R^2$=0.92)**

**Table 3 - Results of Regression Analysis**

| | Coefficient for advice | Coefficient for initial estimate | Overall significance of model | Adj. $R^2$ |
|---|---|---|---|---|
| Extreme Adjustment | $t_{210}$ = -7.68 | $t_{210}$= 14.99 | $F_{1,211}$ = 141.26 | 0.57 |
| | p<0.001 | p<0.001 | p<0.001 | |
| Bracketing | $t_{80}$ = 18.21 | $t_{80}$ = -8.49 | $F_{1,81}$ = 260.72 | 0.87 |
| | p<0.001 | p<0.001 | p<0.001 | |
| Moderate Adjustment | $t_{623}$ = 19.54 | $t_{623}$ = 52.89 | $F_{1,624}$ = 3517.20 | 0.92 |
| | p<0.001 | p<0.001 | p<0.001 | |

Table 3 indicates that all models are significantly meaningful in explaining the variation in final estimates. For moderate adjustment case, the regression model shows, that the final decision, overall, was taken by 1/5* initial forecast + 4/5* advice: that is participants approximately gave four times greater weight to the advice than their initial estimates. Extreme adjustment of advice leads to the negative coefficient of initial estimates. Furthermore in bracketing case, judges negatively use the advice which results in a negative coefficient for the advice.

As it is mentioned before, usefulness index and dominance weights are the measures that compare the model before and after adding a new advice in the form of percentage increase in the criterion variance explained. The difference between the two measures is that, in dominance weight, all the subsets belonging to advices are considered. However, owing to the fact that this study has only one advice, both measures will produce the same results. Therefore, only usefulness index will be used hereafter.

**Table 4 - Comparison of Regressions**

| | $R^2$(single) | $R^2$(multiple) | % change |
|---|---|---|---|
| Extreme Adjustment | 0.12 | 0.57 | 375 |
| Bracketing | 0.75 | 0.87 | 16 |
| Moderate Adjustment | 0.55 | 0.92 | 67 |

In order to calculate the usefulness index, I first regress final estimates on the initial estimates (designated 'single' in Table 4) and then use the previous part's models (designated 'multiple' in Table 4) for the three cases. Table 4 confirms that extreme adjusters totally switch to advisor's recommendation, so they utilize the advice to a great extent. However, moving away from

advice leads to less utilization of advice as shown in the bracketing case. Judges that move closer to advice, utilize 67% from advice according to usefulness index.

### 4.4 Direct Rating Measures

As mentioned before, direct rating is a widely used approach for measuring level of different characteristics or capabilities. In this study, I used 7-point likert scale as Dawes (2008) suggests and the decision makers rated the extent they believe to have utilized the advice for each decision. The correspondence table is seen below;

1 = Never Utilized

2 = Somewhat Utilized

3 = Little Utilized

4 = Moderately Utilized

5 = Very utilized

6 = Often Utilized

7 = Completely Utilized

**Table 5 - Comparison of Means**

| | μ | % |
|---|---|---|
| Extreme Adjustment | 5.6 | 80 |
| Bracketing | 1.3 | 18.6 |
| Moderate Adjustment | 3.55 | 50.7 |

The ratings for the three cases are separated and the mean value is calculated as shown above. Results indicate that people utilized more advice when they move beyond the advice while initial estimate is not supported

after the advice is shown. Nevertheless, people moderately used advice ($\mu$=3.99) in extreme adjustment case. On the other hand, people tend to utilize almost 50% of the advice when they decide on utilizing both advice and their initial estimate. In other words, people give equal weights to their judgment and advisor's recommendation. In case of bracketing the advice, 1.3 mean score indicates that people almost never utilized advice since it causes to move further away.

### 4.5 Comparison of the Utilization Measures

In general, results of the formula-based, regression-based and direct rating techniques reinforce one another. First of all, when judges extremely adjust the advice, mean "Advice Taking" (2.5) indicates that, on average, judges discount their estimates and move more than twice beyond the advice. That is why regression model gives negative weight for the initial estimate while the weight of advice is more than 1 (1.12). Usefulness index also indicates that extreme adjusters utilized the given advice by a factor of 375%. Finally, direct rating measures support that people often utilized advice more than their estimates.

**Table 6 - Comparison of Approaches**

| | $\mu$ (Advice Taking) | $\beta$'s | Usefulness Index | Direct Rating |
|---|---|---|---|---|
| Extreme Adjustment | 250% | 112% | 375% | 80% |
| Bracketing | -180% | -70% | 16% | 19% |
| Moderate Adjustment | 70% | 78% | 67% | 50% |

In the bracketing case, judges are negatively affected by the advice and they move further away from it. An average "Advice Taking" of 1.8 shows that judges move away from the advice almost as twice as the difference between their initial estimate and the advice. For the same reason, regression model indicates positive weight for initial estimate and negative weight for the advice. Besides, usefulness index supports this argument by indicating that only 16% of the advice is utilized. Very similar result is obtained in direct rating approach showing that people never or rarely tend to utilize advice in bracketing case.

As majority of the judges adjust towards the advice to some extent, they consider both their initial estimates and the recommendation of the advice. This result in an average "Advice Taking" of 0.69 and that entails judges give 69% weight to the advice and 31% weight to their initial estimate. In the regression technique, these values change to 78% for the advice and 21% for the initial estimate. Usefulness index is also along similar lines since it claims that 67% of the advice is utilized in the moderate adjustment case. Finally, direct rating results show that people evenly utilized advice and their estimates (51%).

### 4.6 Decision Performance for the Cases

The advice utilization results indicate that judges utilized 112% of the advice in extreme adjustment; -74% of the advice in bracketing and 78% of the advice in moderate adjustment when considering regression-based results. Whether over-utilized, over-discounted and moderately utilized advice help judges to improve decision performance is an important issue. In order to find out that, two accuracy measures will be used. MAE-to-Mean ratio is

selected as the most appropriate measure considering the percentage errors. For the scale errors, MAE is preferred, because it is easier to see how big error is obtained in estimations. In fact, MdAE could also be used since it trims outliers by providing a unit-free measure, but the data interval is so narrow that there is no significant deviation in data range, so the data do not suffer from any outlier. Decision accuracy for each case is calculated in terms of MAE and MAE-to-Mean scores as shown in Table 7 and 8 respectively.

**Table 7 – MAE Results**

| | MAE initial | MAE final | % Change |
|---|---|---|---|
| Extreme Adjustment | 15.63 | 12.82 | -18% |
| Bracketing | 11.29 | 13.77 | 22% |
| Moderate Adjustment | 13.95 | 10.45 | -25% |

**Table 8 – MAE-to-Mean Results**

| | MAE-to-Mean Initial | MAE-to-Mean Final | % Change |
|---|---|---|---|
| Extreme Adjustment | 26.68% | 20.98% | -21% |
| Bracketing | 17.89% | 21.27% | 19% |
| Moderate Adjustment | 20.88% | 16.94% | -19% |

I adjusted the mean values for the three cases by decreasing 1900. The reason is that the mean absolute errors are around 15 and mean values are approximately 1960-1965. This causes MAE-to-Mean ratio to give same results with MAE in terms of % change in accuracy. However, all estimates are within 1900 and 1999. Thus, this exclusion leads to observe same scale for the errors and the mean values. As an illustration, mean value of the true estimates for bracketing case was 1963 and it becomes 63 whereas MAE is

14. Thanks to this adjustment, MAE-to-Mean results differed from MAE results. Nevertheless, both measures give similar results in initial and final estimations except % change.

Table 7 represents mean absolute error results of three cases for the initial and final estimates as well as the % changes. Note that the less MAE, the more accuracy and negative % change means an improvement in decision performance. First of all, extreme adjusters performed badly in terms of initial accuracy whereas judges in bracketing case outperform. This explains the main reason of behaving unusual in decisions. Judges that have sufficient information about a decision task feel high levels of confidence and this causes over-discounted advice. At the same direction, extreme adjusters over-utilize advice, because they are not confident about their initial decisions. This is clearly seen in their initial performance.

Considering the accuracy of final estimates, extreme and moderate adjusters improve decision performance by approximately 20% whereas judges that moves far away from advice performed almost 20% worse. This indicates that use of advice helps judges to increase decision performance (Yaniv, 2004b). Another result is that although MAE results show that moderate adjusters improve more than extreme adjusters, whereas MAE-to-Mean result indicates the opposite. Since there is no significant difference between the improvement degrees as shown in Table 9, it does not cause any conflict in results. In both case, advice is utilized and so final estimates stand close to advice leading to success in accuracy.

As mentioned before, two benchmark models are used to compare with the decision accuracy of judges of three cases. These models are taken from

Yates (1994)'s study about accuracy analysis. The first one is a constant judge whose estimations do not differ with respect to decision tasks. Since a constant estimate for all decision tasks cannot be linked to an advice, it is selected by taking average of 20 true estimates (1957.5) in order to minimize the errors. The second benchmark is uniformly formed by using advice and the initial estimate. It is equally weighted average of them and it means that advice is half utilized and half discounted. This benchmark is expected to perform better than the constant one since both judge's estimate and advice is used in decision. Table 9 shows the t-values and p-values for the comparison of difference between MAE values of initial estimate, final estimate, averaging and constant benchmarks for the three cases. Results indicate that all mean absolute errors are significantly different than each other. Thus, it can be said that these estimations are comparable and they have different accuracy performance.

**Table 9 – Significance Test Results**

| | MAE Initial vs. MAE Final | MAE Final vs. Averaging Benchmark | MAE Initial vs. Averaging Benchmark | MAE Final vs. Constant Benchmark |
|---|---|---|---|---|
| Extreme Adjustment | $t_{210} = 7.34$ | $t_{210} = 8.59$ | $t_{210} = 31.43$ | $t_{210} = 14.62$ |
| | $p<0.001$ | $p<0.001$ | $p<0.001$ | $p<0.001$ |
| Bracketing | $t_{210} = 18.26$ | $t_{210} = 12.85$ | $t_{210} = 26.68$ | $t_{210} = 5.99$ |
| | $p<0.001$ | $p<0.001$ | $p<0.001$ | $p<0.001$ |
| Moderate Adjustment | $t_{210} = 18.64$ | $t_{210} = 28.97$ | $t_{210} = 26.33$ | $t_{210} = 17.08$ |
| | $p<0.001$ | $p<0.001$ | $p<0.001$ | $p<0.001$ |

Results are shown in Table 10 and Table 11 including judge's initial and final performance. Initially, expectation about the constant estimation comes true in a sense that final estimates performed twice better than the constant model. No great difference is obtained in constant model considering the

three cases which is also expected since the characteristics of data are similar. Considering the second benchmark, MAE results show that extreme adjusters performed worse than the judges that select averaging, but MAE-to-Mean ratios give almost same values. This is due to convergence of estimates to the advice for both averaging and extreme adjustment cases. On the other hand, moderate adjusters display higher decision performance than the judges that utilize averaging. This is thanks to the judgment strength of humans including the collaboration between two parties.

**Table 10 – MAE Comparison with Benchmarks**

| | MAE Initial | MAE Final | MAE Averaging Benchmark | MAE Constant Benchmark |
|---|---|---|---|---|
| Extreme Adj. | 15.63 | 12.82 | 12.27 | 22.60 |
| Bracketing | 11.29 | 13.77 | 11.40 | 23.76 |
| Moderate Adj. | 13.95 | 10.45 | 12.18 | 22.04 |

**Table 11 – MAE-to-Mean Comparison with Benchmarks**

| | MAE-to-Mean Initial | MAE-to-Mean Final | MAE-to-Mean Averaging Benchmark | MAE-to-Mean Constant Benchmark |
|---|---|---|---|---|
| Extreme Adjustment | 26.68% | 20.98% | 21.02% | 39.31% |
| Bracketing | 17.89% | 21.27% | 18.51% | 41.32% |
| Moderate Adjustment | 20.88% | 16.94% | 19.57% | 38.33% |

MAE and MAE-to-Mean results for the bracketing case are not surprising. If those judges used averaging technique, then they would outperform to the other cases. That is because when final estimate and advice falls in the

opposite site of the initial estimate (as in the bracketing case), comparing to the final estimate, the average of advice and initial estimate stands closer to the advice and so the true estimate (assuming that it is an expert advice). This result is also supported by Soll and Larrick (2009) which found that averaging achieved much greater accuracy than the case of bracketing when compared to the usual case. Another observation at this point is that initial performance in bracketing case is almost same with the averaging case, but then they get worse after faced with the advice. This is because moving away from an expert advice cause people to reduce decision performance.

# CHAPTER 5

# DISCUSSION AND FURTHER RESEARCH

In decision making literature, different criteria have been developed to measure how individuals use advice for their decisions. The assessment and utilization of advice depends on several factors such as existence of pre-advice estimate, number of advice and the interaction between judge and the advisors. In this thesis, in order to measure the degree of advice utilization, three main methods (one formula-based, one regression based, one direct rating) are compared in an empirical setting in which judges face with expert advice which is randomly generated. Then decision accuracies are analyzed and judges are compared to benchmark models.

When facing with an advice, it is critical for the decision makers to move toward the advice or their initial decision. To investigate this issue, different situations are categorized and three cases are established with respect to adjustment ordering. The novel reason to do that is to separate the "out of range" or 'problematic' instances that judges give negative weight or extreme weight to an advice. This was Bonaccio and Dalal (2006)'s call to focus on these issues. Separating the estimates with respect to the judge's adjustment choice is useful because these three cases represent different dynamics with respect to the utilization of advice.

Advice utilization results confirm that all formula-based, regression-based and direct rating techniques attain results in the same direction.

Nevertheless, regression models explain the interaction between the judge and the advice in a more detailed and informative manner. It provides the degree of variation in the final estimate as explained by the presented advice and the judge's initial estimate, which is not a case for formula-based approaches. On the other hand, formula-based approaches can be used to measure the utilization of a one-time decision, but this simply is not possible with the regression-based approaches. In case of a series of decisions by several judges, the regression method seems the preferable choice when compared against formula-based calculations.

An important issue about formula-based methods is that Yaniv (2004b) took the absolute value of Harvey and Fischer (1997)'s formula and presented WOA just in order to eliminate the negative values formed by only the bracketing case. This manipulates the mean of a data which indicates the degree of advice utilization, because advice is over-discounted in bracketing case and that is why "advice taking" ratio becomes negative. Taking the absolute value of that ratio contributes to the mean as over-utilized advice, but that is not the case. In short, it gives unhealthier results to use WOA in decisions that include bracketing case. Although that case rarely happens, it does not mean that it should be ignored. For the future research, scholars that utilize from WOA should consider the "out of range" decisions.

The third utilization method, direct rating approach is taking simply the average of ratings to indicate the degree of advice utilization, but it is obvious that results should be similar to the other utilization measures unless people are dishonest in explaining the truth. At this point, results of direct rating technique show a different aspect of decision making framework. When people utilized both advice and the initial estimate, they

announce less utilization ratio than they actually utilized advice. This may be due to the psychological factors such as egocentric behavior presented by Yaniv & Kleinberger (2000). This kind of a behavior is the result of tendency of judges to underweight advice. Parallel to this argument, people may tend to avoid announcing that they use advice a lot. The question that needs to be answered is whether people actually avoid announcing it or they are not aware of the degree of advice utilization. Further study should contain the analysis of this tendency in decision makers.

The novel motivation behind advice literature is to achieve higher accuracy in decisions. The whole team and group decision literatures aim to achieve better results in decision performance of individuals and institutions. Advice taking plays an important role in decision analysis in a sense that the extent of improvement in decision after receiving and especially reacting to an advice. At this point, the way of reaction plays an important role in improving decision performance. Two accuracy measures, MAE and MAE-to-Mean are used and they exhibit similar results due to indifference between estimation mean of cases. However, the three cases yield different results in terms of accuracy. Judges that bracket their estimate get worse comparing to the initial estimates whereas the other two groups enhanced their decisions. All cases outperform than a constant judge, but averaging seems to surpass the bracketing method.

An interesting observation of accuracy analysis is that initial performance in bracketing case is almost same with the averaging case. Cooper (1991) argued that confident people underutilize advice than less confident ones. Then, judges that bracket their estimates show more confidence which is reflected in their decision and performance since it is a sign of knowledge or

accuracy level about the decision tasks. Looking from the opposite perspective, Chaiken, Liberman and Eagly (1989) found that low confidence in accuracy can lead to solicit additional information. The critical question is whether lack of confidence in accuracy can be a sign of low accuracy.

Previous research on this issue shows that there is not strong but a significant relationship between confidence and accuracy (e.g., Deffenbacher, 1980; Sporer, Penrod, Read & Cutler, 1995). Therefore, it should be more widely investigated for the future research that confident people seems to perform a little better than unconfident ones, but they over-discount additional opinions and so they are defeated to the unconfident people who solicit expert advice. Further studies should involve confidence ratings of judges in their initial estimates so that these issues can be investigated. Another important issue is that the power of random advice is not deeply analyzed by the JAS scholars. The motivation behind the usage of random advice and its representativeness in form of non-expert advice should also be investigated in a detail manner for the future research.

There were some limitations about experimental design and methodology. The credibility of the source of advice is not declared to the judges in detail. However, this can be eliminated by including feedback about an advisor's decision performance during each decision task in studies that focus on the representativeness of random advice for the future research. Another limitation was that the number subjects can be more so that the sample size could enable to get more information for the three cases. Further studies should include larger sample size to more reliably reflect an unusual behavior. To sum up, research on adjustment behavior of decision makers represents a new frontier for advice research.

# REFERENCES


Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: *Prentice-Hall*.

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8, 69-80.

Asch, S. E. (1952). Social psychology. Oxford, United Kingdom: Prentice Hall.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, 8, 129–148.

Baillon, Aurélien, Laure Cabantous, & Peter P. Wakker (2012), "Aggregating Imprecise or Conflicting Beliefs: An Experimental Investigation using Modern Ambiguity Theories," *Journal of Risk and Uncertainty*, vol. 44, pp. 115–147.

Basow, S. A., & Rubenfeld, K. (2003). "Troubles talk": Effects of gender and gender-typing. *Sex Roles*, 48, 183–187.

Billings, R. S., & Scherer, L. L. (1988). The effects of response mode and importance on decision-making strategies: Judgment versus choice. *Organizational Behavior & Human Decision Processes*, 41, 1-19.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, 37, 48–74.

Björnsson, Y., Hafsteinsson, V., Jóhannsson, A., & Jónsson, E. (2004). Efficient Use of Reinforcement Learning in a Computer Game, *in Proceedings of the International Conference on Computer Games:* Artificial Intelligence, Design and Education, pp. 379-383.

Bonaccio, S., & Dalal, R. S. 2006. Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, Vol. 101*, pp. 127-151.

Brehmer, B., & Hagafors, R. (1986). The use of experts in complex decision-making: a paradigm for the study of staff work. *Organizational Behavior and Human Decision Processes*, 38, 181–195.

Brunswik, E. (1955). Representative design and probabilistic theory in functional psychology. *Psychological Review*, 62, 193–217.

Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.). Berkeley, CA: *University of California Press*.

Budescu, D. V. 2006. Confidence in aggregation of opinions from multiple sources. K. Fiedler, P. Juslin, eds. Information Sampling and Adaptive Cognition. *Cambridge University Press*, New York, 327–352.

Budescu, D. V., & Rantilla, A. K. (2000). Confidence in aggregation of expert opinions. *Acta Psychologica*, 104, 371–398.

Budescu, D. V., Rantilla, A. K., Yu, H., & Karelitz, T. K. (2003). The effects of asymmetry among advisors on the aggregation of their opinions. *Organizational Behavior and Human Decision Processes*, 90, 178–194.

Budescu, D., & Yu, H.-T. (2007). Aggregation of opinions based on correlated cues and advisors. J. Behav. *Decision Making*, 20, 153-177.

Carbone, R., & Armstrong, J. S. (1982). Evaluation of extrapolative forecasting methods: Results of a survey of academicians and practitioners. *Journal of Forecasting*, 1, 215 – 217.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), Unintended thought (212–252). New York: Guilford.

Clarebout, G., & Elen, J. (2008). Advice on tool use in open learning environments. *Journal of Educational Multimedia and Hypermedia*, 17(1), 81-97.

Cohen, P.R., Oates, T., & St. Amant, R. (1996). Plan Steering and Mixed-Initiative Planning. In A. Tate, ed. ARPI Supplement to Proc. 3rd Intl. Conf. on AI Planning Systems, 105-112.

Cooper, R. S. (1991). Information processing in the judge–adviser system of group decision-making. Unpublished master's thesis, University of Illinois, Urbana-Champaign.

Dalal, R. S. (2001). The effect of expert advice and financial incentives on cooperation. *Unpublished master's thesis*, University of Illinois, Urbana-Champaign.

Dalal, R. S., & Bonaccio, S. (2010). What types of advice do decision-makers prefer? *Organizational Behavior and Human Decision Processes*, 112(1), 11-23.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161–182.

Dawes, J. (2008) "Do data characteristics change according to the number of scale points used?" *International Journal of Market Research*, Vol 50, No. 1, 61-77.

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason, (Ed.), Test anxiety: Theory, research, and applications, 111–124. Hillsdale, NJ: Erlbaum.

Einhorn, H.J., R.M. Hogarth & E. Klempner, (1977). Quality of Group Judgement. *Psychological Bulletin*. 84, 158-172.

Feng, B., & MacGeorge, E. L. (2006). Predicting receptiveness to advice: Characteristics of the problem, the advice-giver, and the recipient. *Southern Communication Journal*, 71, 67–85.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.

Fischer, I., & Harvey, N. (1999). Combining forecasts: what information do judges need to outperform the simple average? *International Journal of Forecasting*, 15, 227–246.

French, J. R. P., Jr., & Raven, B. H. (1959). The bases of social power. In D. Cartwright (Ed.), Studies in social power (pp. 150–167). Ann Arbor, MI: *Institute for Social Research*.

G. Gigerenzer, P. M. Todd, & The ABC Research Group. Simple heuristics that make us smart. *Oxford University Press*, Oxford, UK, 1999.

Gardner, D. H. & Berry, D. C. (1995). The effect of different forms of advice on the control of a simulated complex system. *Applied Cognitive Psychology*, 9, 555–579.

Gibbons, A. M. (2003). Alternative forms of advice in natural decision settings. Unpublished master's thesis, University of Illinois,Urbana-Champaign.

Gibbons, A. M., Sniezek, J. A., & Dalal, R. S. (2003, November). Antecedents and consequences of unsolicited versus explicitly solicited advice. In D. Budescu (Chair), Symposium in Honor of Janet Sniezek. Symposium presented at the *annual meeting of the Society for Judgment and Decision Making*, Vancouver, BC.

Goldsmith, D. J., & Fitch, K. (1997). The normative context of advice as social support. *Human Communication Research*, 23, 454–476.

Handgraaf, M. J. J., Milch, K. F., Appelt, K. C., Schuette, P., Yoskowitz, N. A., & Weber, E. U. (2012). Web-conferencing as a viable method for group decision research. *Judgment and Decision Making*, 7(5), 659-668.

Hardman, D. (2009). Judgment and decision making: Psychological perspectives. Chichester: BPS Blackwell.

Harvey, N., & Fischer, I. (1997). Taking advice: accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, 70, 117–133.

Harvey, N., Harries, C., & Fischer, I. (2000). Using advice and assessing its quality. *Organizational Behavior and Human Decision Processes*, 81, 252–273.

Hawitt, D. (2010), Should you report forecast error or forecast accuracy, *Foresight*, 18, 46.

Hedlund, J., Ilgen, D. R., & Hollenbeck, J. R. (1998). Decision accuracy in computer-mediated versus face-to-face decision-making teams. *Organizational Behavior and Human Decision Processes*, 76, 30–47.

Hill, G. W. (1982). Group Versus Individual Performance : Are TV + 1 Heads Better Than One ?, 91(3), 517–539.

Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., Hedlund, J., Major, D. A., & Phillips, J. (1995). Multilevel theory of team decision making: decision performance in teams incorporating distributed expertise. *Journal of Applied Psychology*, 80, 292–316.

Hoover, J. (2006). Measuring forecast accuracy: Omissions in today's forecasting engines and demand planning software, Foresight: *The International Journal of Applied Forecasting*, 4, 32-35.

Hoover, J. (2009). How to track forecast accuracy to guide forecast process improvement. *Foresight*, 14, 17-23.

Hoover, J., & Little, M. (2009). How to track forecast accuracy to guide forecast process improvement. Foresight: *The International Journal of Applied Forecasting*, 2009(14), 17-23.

Humphrey, S. E., Hollenbeck, J. R., Meyer, C. J., & Ilgen, D. R. (2002). Hierarchical team decision making. In G. R. Ferris & J. J. Martocchio (Eds.). *Research in personnel and human resources management*, Vol. 21, pp. 175–213, Stamford, CT: JAI Press.

Hyndman, R. J. & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22:679–688.

Jungermann, H. & Fischer, K. 2005. Using Expertise and Experience for Giving and Taking Advice. In: T. Betsch and S. Haberstroh, eds., *The Routines of Decision Making*. Taylor & Francis.

K. R. Smith, (1999) Using the Same Decision making Process for Joint and Army Operations. *School of Advanced Military Studies.*

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. New York: *Cambridge University Press*.

Kennedy, J., D. N. Kleinmuntz, & M. E. Peecher. 1997. Determinants of the justifiability of performance in ill-structured audit tasks. *Journal of Accounting Research* 35 (Supplement): 105-123.

Kolassa, S. & Schütz, W. (2007). Advantages of the MAD/MEAN ratio over the MAPE, Foresight: *The International Journal of Applied Forecasting*, Issue 6, 40-43.

Kruger, J., Wirtz, D., & Miller, D. T. (2005). Counterfactual thinking and the first instinct fallacy. *Journal of Personality and Social Psychology*, 88, 725-735.

Lim, J. S., & O'Connor, M. (1995). Judgmental adjustment of initial forecasts: its effectiveness and biases. *Journal of Behavioral Decision Making*, 8, 149–168.

Makridakis, S., & Wheelwright, S.C. (1989), *Forecasting Methods for Management* (5th edition). Chichester: Wiley.

Makridakis, S., S. C. Wheelwright, & V. E. McGee. (1983). *Forecasting: Methods and Applications*. New York: John Wiley and Sons.

March JG, Simon HA. 1958. *Organizations*. New York: Wiley.

Meshi, D., Biele, G., Korn, C. & Heekeren, H. 2012, "How Expert Advice Influences Decision", *PLoS ONE*, [Online], , pp. 21.

Messick, D. M., & Ohme, R. K. 1988. Some ethical aspects of the social psychology of social influence. In R. M. Kramer & M. A. Neale (Eds.), Power and influence in organizations: 181-202. Thousand Oaks, *CA: Sage Publications*, Inc.

Önkal, D., P. Goodwin, M. Thomson, M.S. Gönül, and A.Pollock (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, 22, 390–409.

Oxford Dictionary and Thesaurus (1997) Oxford, UK, p.24.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. Cambridge, MA: *Cambridge University Press.*

Phillips, J.M. (1999). Antecedents of leader utilization of staff input in decision-making teams. *Organizational Behavior and Human Decision Processes*, 77, 215-242.

Price, P. C., & Stone, E. R. (2004). Intuitive evaluation of likelihood judgment producers: *Evidence for a confidence heuristic*. Journal of Behavioral Decision Making, 17(1), 39–57.

Raven, B. H. (1965). Social influence and power. In I.D. Steiner & M. Fishbein (Eds.), *Current studies in social psychology* (pp. 371–382). New York: Holt, Rinehart, Winston.

Ray, W. S. (1955). Complex tasks for use in human problem-solving research. *Psychological Bulletin*, 52, 134–149.

Savadori, L., Van Swol, L. M., & Sniezek, J. A. (2001). Information sampling and confidence within groups and judge advisor systems. *Communication Research*, 28, 737–771.

Schotter, A. (2003). Decision making in the face of naive advice. *American Economic Review Papers & Proceedings*, 93(2), 196-201.

Schrah, G. E. (2000). Task complexity and strategy selection in Judge–Advisor decision making: A cost–benefit model of advice utilization. Unpublished master's thesis, University of Illinois, Urbana-Champaign.

Schrah, G. E., Dalal, R. S., & Sniezek, J. A.(2006). The adaptive decision maker is not an island: integrative expert advice with information search. *Journal of Behavioral Decision-Making*, 19, 43–60.

Simon, H. A. (1982). Models of bounded rationality (2 vols.). Cambridge, MA: *MIT Press*.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge–advisor decision making. *Organizational Behavior and Human Decision Processes*, 62, 159–174.

Sniezek, J. A., & Henry, R. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, 43(1), 1-28.

Sniezek, J. A., & Henry, R. (1990). Revision, weighting, and commitment in consensus group judgment. *Organizational Behavior and Human Decision Processes*, 45(1), 66-84.

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge–advisor system. *Organizational Behavior and Human Decision Processes*, 84, 288–307.

Sniezek, J. A., Schrah, G. E., & Dalal, R. S. (2004). Improving judgment with prepaid expert advice. *Journal of Behavioral Decision Making*, 17 , 173–190.

Soll, J. B., & Larrick, R. P. (2004). Strategies for revising judgment: How, and how well, do people use others' opinions? Unpublished manuscript, Duke University.

Soll, J.B., & Larrick, R. P. (2009). Strategies for revising judgment: how (and how well) people use others' opinions, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35, 780–805.

Sporer, S. L., Penrod, S. D., Read, J. D., & Cutler, B. L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin*, 118, 315–327.

Sundstrom, E., De Meuse, K. P., Futrell, D. (1990) Work teams: applications and effectiveness, *American Psychologist*, 120 – 133.

Surowiecki, J., The Wisdom of Crowds, New York: Doubleday, 2004; *Anchor Books*, 2005.

Swanson, D., J. Tayman, & T. Bryan. (2011). MAPE-R: A rescaled measure of accuracy for cross-sectional, sub-national forecasts. *Journal of Population Research* 28, 225-243.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Van Swol, L. M. (2009). The effects of confidence and advisor motives on advice utilization. *Communication Research*, 36(6), 857–873.

Van Swol, L. M., & Ludutsky, C. L. (2007). Tell me something I don't know: Decision makers' preference for advisors with unshared information. *Communication Research*, 34, 297-312.

Van Swol, L. M., & Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology*, 44, 443–461.

Vroom, V. H. & Philip W. Y. (1973). Leadership and Decision-making. Pittsburgh, PA: *University of Pittsburgh Press*,

Yaniv, I. (1997). Weighting and trimming: heuristics for aggregation of judgments under uncertainty. *Organizational Behavior and Human Decision Processes*, 69, 237–249.

Yaniv, I. (2004a). The Benefit of additional opinions. *Current Directions in Psychological Science,* 13, 75–78.

Yaniv, I. (2004b). Receiving other people's advice: influence and benefit. *Organizational Behavior and Human Decision Processes*, 93, 1–13.

Yaniv, I., & Kleinberger, E. (2000).Advice taking in decision making: egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, 83, 260–281.

Yaniv, I., Choshen-Hillel, S. & Milyavsky, M. (2009). Spurious consensus and opinion revision: Why might people be more confident in their less accurate judgments? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 558-563.

Yates, J. F. 1994. Subjective probability accuracy analysis. in Wright, G. and Ayton, P. (eds), *Subjective Probability*, Chichester: Wiley.

Yates, J. F., Lee, J. and Shinotsuka, H. (1996). Beliefs about overconfidence, including its cross-national variation. *Organizational Behavior and Human Decision Processes*, 65, 138-47.

Zander, A. (1979). The psychology of group research. *Annual Review of Psychology*, 30, 417–451.

# APPENDICES

## APPENDIX A: QUESTIONNAIRE FORM

adsheet/viewform?formkey=dGFJazM4Z3BFd3QtdTNSalEzdHlPVVE6MQ#gid=0

### 20. YÜZYIL İCADLARI- 1. AŞAMA (ADVICE TAKING IN DECISION MAKING)

Aşağıda 20. yüzyıla ait 20 adet buluşun icad tarihlerini yıl olarak tahmin edeceksiniz. Tahminleriniz 1900-2000 arası olmalıdır. Tahmin ederken bağlantılı icadların tarihleriyle, yaygınlaşma tarihleriyle ya da çeşitli olaylarla ilişkilendirebilirsiniz. Bu anket 1. aşamadır, 2. aşamada size genel kültür seviyesi iyi birinin tahminleri verilecek ve son kararınızın verilmesi istenecektir. Lütfen 1. anketi doldurduktan sonra gönder'e TIKLAMAYINIZ ve sayfayı kapatmayınız çünkü son kararınız için ilk kararınıza bakmak isteyebilirsiniz. Tahminlerinizi girdikten sonra 2. aşamaya geçiniz. (Lütfen tahminde bulunurken bu icadların evrensel olduğunu ve ülkemizin teknolojik açıdan avrupayı ne yazık ki geriden takip ettiğini aklınızda bulundurunuz)

* Gerekli

**1. İnternet kaç yılında icad edilmiştir?** *

**2. ATM makinesi kaç yılında icad edilmiştir?** *

**3. Mikrodalga fırın kaç yılında icad edilmiştir?** *

**4. Radyo kaç yılında icad edilmiştir?** *

**5. Penisilin kaç yılında bulunmuştur?** *

**6. Uydu ilk kez Dünyanın yörüngesine kaç yılında yerleştirilmiştir?** *

**7. Renkli TV kaç yılında icad edilmiştir?** *

8. İlk kalp nakli kaç yılında gerçekleştirilmiştir? *

9. Radar kaç yılında icad edilmiştir? *

10. İlk tüp bebek kaç yılında icad edilmiştir? *

11. Kredi kartı kaç yılında icad edilmiştir? *

12. Scanner (tarayıcı) kaç yılında icad edilmiştir? *

13. Kol saati kaç yılında icad edilmiştir? *

14.Motorlu araba (benzinle çalışan) kaç yılında icad edilmiştir? *

15.CD (Compact Disc) kaç yılında icad edilmiştir? *

16. Navigasyon cihazı kaç yılında icad edilmiştir? *

17. Cep telefonu kaç yılında icad edilmiştir? *

18. MP3 çalar kaç yılında icad edilmiştir? *

19. İlk E-mail kaç yılında atılmıştır? *

20. Google kaç yılında bulunmuştur? *

adsheet/viewform?formkey=dDI5SnEwYnJ3ZE9iNWgzOUZYMmI2RXc6MQ#gid=0

# 20. YÜZYIL İCADLARI- 2. AŞAMA (ADVICE TAKING IN DECISION MAKING)

1. aşamada 20. yüzyıla ait 20 adet buluşun icad tarihlerini yıl olarak tahmin ettiniz. Bu aşamada ise size genel kültür seviyesi iyi birinin tahminleri "Advice" olarak verilecek ve son kararınızın verilmesi istenecektir. Lütfen 1. anketteki ilk kararınıza bakarak size verilen tahminleri de göz önünde bulundurarak son kararınızı veriniz. İki anket için de GÖNDER'e basmayı unutmayınız. Bu çalışmanın barındırdığı teorileri veya metodolojiyi öğrenmek için ali.ozarslan@metu.edu.tr'ye mail atabilirsiniz. Yardımcı olduğunuz için çok teşekkürler!!!

* Gerekli

**1. İnternet kaç yılında icad edilmiştir?** *

Advice: 1978 İlk kararınız:

**1. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**2. ATM makinesi kaç yılında icad edilmiştir?** *

Advice: 1970 İlk kararınız:

**2. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**3. Mikrodalga fırın kaç yılında icad edilmiştir?** *

Advice: 1952 İlk kararınız:

**3. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**4. Radyo kaç yılında icad edilmiştir?** *

Advice: 1922 İlk kararınız:

**4. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**5. Penisilin kaç yılında bulunmuştur?** *

Advice: 1909 İlk kararınız:

**5. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**6. Uydu ilk kez Dünyanın yörüngesine kaç yılında yerleştirilmiştir?** *

Advice: 1943 İlk kararınız:

**6. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**7. Renkli TV kaç yılında icad edilmiştir?** *

Advice: 1959 İlk kararınız:

**7. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**8. İlk kalp nakli kaç yılında gerçekleştirilmiştir?** *

Advice: 1979 İlk kararınız:

**8. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**9. Radar kaç yılında icad edilmiştir?** *

Advice: 1916 İlk kararınız:

**9. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**10. İlk tüp bebek kaç yılında icad edilmiştir?** *

Advice: 1962 İlk kararınız:

**10. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**11. Kredi kartı kaç yılında icad edilmiştir?** *

Advice: 1955 İlk kararınız:

**11. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**12. Scanner (tarayıcı) kaç yılında icad edilmiştir?** *

Advice: 1962 İlk kararınız:

**12. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**13. Kol saati kaç yılında icad edilmiştir?** *

Advice: 1924 İlk kararınız:

**13. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**14.Motorlu araba (benzinle çalışan) kaç yılında icad edilmiştir?** *

Advice: 1918 İlk kararınız:

**14. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**15.CD (Compact Disc) kaç yılında icad edilmiştir?** *

Advice: 1977 İlk kararınız:

**15. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**16. Navigasyon cihazı kaç yılında icad edilmiştir?** *

Advice: 1983 İlk kararınız:

**16. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**17. Cep telefonu kaç yılında icad edilmiştir?** *

Advice: 1981 İlk kararınız:

**17. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**18. MP3 çalar kaç yılında icad edilmiştir?** *

Advice: 1992 İlk kararınız:

**18. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**19. İlk E-mail kaç yılında atılmıştır?** *

Advice: 1979 İlk kararınız:

**19. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

**20. Google kaç yılında bulunmuştur?** *

Advice: 1989 İlk kararınız:

**20. soruda "Advice"dan ne kadar yararlandınız?** *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Hiç yararlanmadım | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Çok yararlandım |

Gönder

# APPENDIX B: TURKISH SUMMARY

Gerek üniversite seçimi gibi zor bir karar olsun gerekse bir mekana ulaşmak için yolu bulmak gibi basit bir karar olsun; insanlar çoğu zaman bu gibi durumlarda zorluk yaşarlar. Bu tür kararların ortak özelliği ise ortada belirsizliğin mevcut olmasıdır. Festinger (1954)'e göre karar alırken belirsizlikleri çözmek içsel bir ihtiyaçtır. Neyse ki birçok kez, bu tür karar aşamaları başkalarından yardım alarak daha kolay hale getirilir. Fazladan bir düşünce her zaman insanın kendi düşüncesinin yanında kullanması için iyidir. Nitekim, Solomon Asch (1952)'in araştırmasına göre bireyler görünüşte yanlış olsa bile grup kararına sadık kalmayı tercih etmektedirler. Ayrıca, araştırmalar insanların algıları ve seçimlerinin başkalarının görüşleri etkisinde olduğunu göstermektedir. (Ajzen & Fishbein, 1980).

İnsanlar her zaman her durumda daha doğru kararlar almak isterler. Karar alırken muhakeme yeteneğini kullanmak ve eldeki bilgileri doğru şekilde ölçmek çok önemlidir. Peki gerek kendi muhakeme gücü olsun gerekse etraftan aldığın tavsiyeleri nasıl kullanılacağı çok önemlidir. Bu bağlamda, karar alırken başkalarının görüşleriyle kendi görüşlerimiz harmanlama konusu çok kritiktir. Tavsiye almanın karar sürecindeki rolü daha önce araştırmacılar tarafından birçok açıdan incelenmiştir (Bonaccio & Dalal, 2006). Literatür taraması kısmında çeşitli faktörlere değinilecektir.

Çoğu kararlar karmaşık bir şekilde yapılandırılmıştır. Karmaşıklığı azaltmak için bazı sezgisel ilkeler üzerinde tutunma eğilimi gözlemlenmiştir (Tversky ve Kahneman , 1974). Fakat bu ilkelerin insanı en iyi değil sadece yeterli olan kararlara sürüklemiştir. Bu yüzden bu ilkelere tutunmamak ve muhakeme

gücüne başvurmak tavsiye edilmiştir. Bu da insanların karar alırken danışmanların tavsiyelerini almaya sevketmiştir. Bu tezde, insanların tavsiye kullanma davranışları ve kullanılan tavsiyelerin karar performansını arttırıp arttırmadığı incelenmiştir. Bu bağlamda tavsiye çeşitlerinden rastgele tavsiye kullanılmıştır. Belirli bir hata payı konularak rassal olarak tavsiye edilecek tahminler üretilmiştir. Rassal tavsiye oluşturulduktan sonra anket metoduyla denekleri karar verici pozisyonuna getirilmiştir.

Deneklere çeşitli tahminler yapacakları sorular sorulmuş ve tavsiye verimiştir. Daha sonra tavsiyelerden ne kadar yararlandığı ve ne kadar karar performansını arttırdıkları incelenmiştir. Burada literatürdekilerden farklı bir metod izlenmiş ve tavsiyeye yaklaşan insanlar, tavsiyeden uzaklaşan insanlar ve tavsiyeyi aşan insanlar olmak üzere 3 gruba ayrılmıştır. Bu 3 grubun herbiri için tavsiyeden yararlanma dereceleri ve karar tutarlılığını arttırma dereceleri ölçülerek kıyas yöntemi kullanılmıştır. Bunları yapmak için formül tabanlı , regresyon tabanlı ve ölçek bazlı tavsiye kullanım ölçüm teknikleri kullanılmıştır. Daha sonra çeşitli tutarlılık ölçütleri kullanılarak performanslar kıyaslanmıştır.

Tavsiye alma literatürüne bakmadan önce karar verme modellerini incelemekte fayda vardır. Öncelikle Brunswick (1955, 1956) tarafından "lens modeli" adı verilen bir karar verme modeli geliştirilmiştir. Lens modeline göre karar verirken ilgili her ipucuna birer olasılık atayarak lineer fonksiyonla gerçekle arasındaki ilişki kurulabilmektedir. Brunswick'in lens modeli ideal bir çerçeve sunarak bireysel karar sürecini özetlemektedir. Fakat gerek günlük hayatta olsun gerekse örgütsel bağlamda olsun kararların çoğu grup bazında verilmektedir. Örgütsel bağlamda gruplar yöneticiler tarafından oluşturulan ve ortak hedefler oluşturan yapılardır.

Çalışmalar takım dinamiklerinin karar performansı üzerinde önemli bir rol oynadığını ortaya koydu . (Sundstrom , De Meuse ve Futrell , 1990 ).
Hollenbeck ve arkadaşları (1995) "hiyerarşik karar verme ekipleri" diye tanımladığı yapıda karar veren partinin genelde takımın lideri olduğunu ve diğer üyelerin kişisel muhakemeleriyle lidere çeşitli tavsiyelerde bulunduğunu göstermektedir. Brehmer ve Hogafors (1986 ) lens modeline hiyerarşik ekibi ekledi ve "personel karar verme modeli"'ni geliştirdi. Bu modelde, her üye farklı ipuçlarına sahiptir ve nihai kararla ilgili tek tek görüş verir. Ekibin lideri, bu önerileri birleştirir; analiz eder ve sonra karar verir. Grupların karar verme literatüründe bireysel ve grup performansı karşılaştırılmıştır. Çalışmaların çoğu problem çözme çerçevesinde bireylerin ve grupların performansına çeşitli faktörlerin etkisini ölçmüşlerdir. Bu faktörler dört ana başlık etrafında toplanabilir; Görev, süreç, bireysel farklılıklar ve yöntem (Hill, 1982). Neredeyse benzer faktörler tavsiye alma ve kullanımı literatüründe de incelenmiştir (Bonaccio & Dalal , 2006).

Tavsiye teriminin sözlükteki tanımı gelecekteki eylem ya da davranışlar hakkında bir görüş veya öneri olarak sunulan kelimelerdir. Gibbons (2003) tavsiye tanımını genişleterek duygusal desteği ve yeni muhakeme ve alternatif yolları önermeyi de tavsiye tanımına eklemiştir. Çeşitli tavsiye türleri vardır. Yardımcı tavsiye insana ne yapması gerektiğini önerirken karşı tavsiye ise ne yapmaması gerektiğini söyler. Destek ve bilgi de diğer tavsiye türleridir. Ne şekilde olursa olsun peki insanlar neden bir sorun ya da konu hakkında karar vermeden önce ilave görüş almaya ihtiyaç duyar? Yates ve arkadaşlarına göre (1996) insanlar bakış açılarını geliştirmek için tavsiye alırlar. Schotter (2003)'e göre sorunlara yeni yollardan bakmak birincil neden olarak gösterilmiştir. Önemli bir motivasyon ise karar performansını arttırma isteğidir (Yaniv, 2004a; 2004b). Son olarak da tavsiye

alarak başkalarını da kara sürecine eklemek amacıyla kararın sonuçlarını ve etkilerinin verdiği sorumluluğu paylaşmak amacıyla insanlar tavsiye almayı tercih etmektedirler (Harvey ve Fischer, 1997).

Sniezek ve arkadaşları yargıç-danışman sistemini (YDS) ilke kez ortaya atan araştırmacılardır (Sniezek ve Henry, 1990 Sniezek ve Henry, 1989). Burada yargıç muhakeme gücünü kullanarak karar veren kişiyi temsil ederken danışman ise tavsiye veren kişiyi temsil etmektedir. YDS sistemi kapsamında araştırmacılar çeşitli faktörlerin etkisine yoğunlaşmışlardır. İlk olarak tavsiye veren danışmanların yargıç gözündeki tanınırlığı ve güvenilirliği konusunda çeşitli çalışmalar yapılmıştır. ( örneğin Feng ve MacGeorge 2006; Goldsmith ve Fitch, 1997). Danışmanların uzmanlığı konusunda bilgi sahibi olmak yargıçların danışmanlara güvenini arttırmaktadır. (Birnbaum ve Stegner , 1979).

Bazı çalışmalar ve kararlar uzman tavsiyesi ile acemi tavsiyenin etkisini karşılaştırmaktadırlar (örneğin, Meshi , Biele , Korn ve Heekeren , 2012; Jungermann ve Fischer , 2005). Özellikle Meshi ve arkadaşları uzman tavsiyesi kullanımı sırasında beyin aktivitesi üzerine konsantre olmuşlardır. Beynin belirli bölgelerinde sinyalleri kontrol ederek , onların uzman tavsiyelere daha önem verdiklerini bulmuşlardır. Ayrıca, Sniezek ve arkadaşları (2004) uzman tavsiyesinin acemi tavsiyeden daha doğru olduğunu ve bu uzman tavsiyesini kullanan yargıçların karar performansının arttığını gözlemlemişlerdir.

Güven konusu YDS literatüründe büyük öneme sahiptir. Sniezek ve Henry (1989) güven ve karar kalitesini kararlarda eşit öneme sahip olduklarını savunmuşlardır. Karşılıklı güvenin oluşabilmesi için karar sürecinde

tarafların birbiriyle etkileşimi gerekir. Danışmanların birden çok olduğu durumlarda danışmanlar arasındaki tutarlılık da yargıcın güvenini etkiler. Danışmanlar arasında diğerlerinden daha uç tahminlerde bulunan danışmanın mevcudiyeti durumunda yargıç bu danışmanı görmezden gelme yolunu tercih eder.

Daha önce belirtildiği gibi, tipik YDS çalışmaları bir danışman - bir yargıç şeklinde gerçekleştirilir (örn., Harvey ve Fischer , 1997). Ancak birçok çalışmada birden çok danışman bulundurarak yargıçların hangi tavsiyeye uyduğunu ve tavsiye kalitesini ölçmeye çalışmışlardır. Yaniv ve Milyavsky (2007) tarafından yapılan bir başka çalışmada, hakimler kendi görüşünü (yaklaşık %40) ağırlık verip danışman sayısından bağımsız olarak kalan kısmı dağıtma yolunu seçmişlerdir. Diğer önemli bir konu ise tavsiyenin insanlar tarafından istenip istenmemesidir. Çalışmalar şunu göstermiştir ki istenmeden alınan tavsiyeler insanlarda ters teperek bu tavsiyeleri kullanma derecesinin oldukça düşük olduğu gözlemlenmiştir. Fakat insanların isteyerek aldığı tavsiyelerde kulanım oranının yüksek olduğu söylenebilir.

Bir diğer konu ise motivasyon çeşidi olarak deney ortamında karar vericilere karar performansına dayalı finansal destek verilip verilmeyeceğidir. Çalışmalar göstermiştir ki finansal destek alan insanlar tavsiyelerin maliyeti sebebiyle daha az yararlanmışlardır. Fakat karar tutarlılıklarının arttığı gözlemlenmiştir. Önemli konulardan biri de tavsiyenin danışmanlardan deney sırasında alınıp alınmadığıdır. Bazı çalışmalar tavsiyeyi önceden danışmanlardan toplamıştır. Bazıları ise hiç danışman kullanmamış ve bilgisayar kullanarak tavsiye olacak tahminleri oluşturmuşlardır. Bilgisayar kullanan araştırmacılar gerçek tahminleri kullanarak belli bir hata payı koyup rassal tavsiye oluşturmuşlardır. Bazı çalışmalar gerçek tahminleri

hata payı koymadan kullanarak tavsiye olarak belirlemişlerdir. Burada önemli olan hata payının boyutudur. Hata payını çok koyarsa uzman olmayayn danışman; az koyarsa da uzman danışman rolünü üstlenebilir. Bu konuların dışında danışmanların ilk görüşlerinin alınması, yargıç ve danışmanlar arasındaki etkileşim, karar verilecek yada tahmin edilecek soruların çeşidi ve zorluğu da süreçle ilgili olan faktörlerdir.

Yukarıda anılan çalışmalar YDS sisteminin çeşitli dinamiklerini araştırmış ve insanların tavsiyeden faydalanma derecesini ölçmek için çeşitli ölçütler kullanmışlardır. Bonaccio ve Dalal (2006)'nın da özetlediği gibi bu çalışmaların ilki formül bazklı ölçütlerin konu alındığı çalışmalardır. İlk olarak Harvey ve Fischer (1997) "tavsiye alma" oranını ortaya atmıştır. Bu oran yargıcın ilk kararı ile son kararını farkının yargıcın ilk kararı ile tavsiye arasındaki farka oranıyla ölçülür. Yani kendi kararından ne kadar saptığının tavsiyeden ne kadar uzak olduğuna oranı olarak da tanımlanabilir. Örneğin yargıcın ilk kararının 50 olduğunu ve tavsiyenin de 100 olduğunu varsayarsak eğer yargıç son karar olarak 60'I seçerse "tavsiye alma" oranı %20 olur. Yani tavsiyenin %20'si ve kendi kararının %80'inin kulanmış demektir. "tavsiye alma oranı negatif veya 1'den büyük olabilir.

Diğer bir yöntem ise Yaniv (2004b) tarafından ortaya atılmıştır. WOA (tavsiyenin ağırlığı) adı verilen ölçütte Harvey ve Fischer (2000)'in formülünün mutlak değerini alarak negatif değerlerin oluşmamasını amaçlanmıştır. Son olarak da Yaniv ve Kleinberger (2000) Yaniv (2004b)'nin formülünü 1'den çıkararak insanların tavsiyeden yararlanmama oranını oluşturmuşlardır. Bu orana WOE (kendi tahmininin ağırlığı) adını vermişlerdir. Daha önceki örnekteki tahminleri hesaba katarsak WOE oranı %80 olarak ortaya çıkar ve bu da tavsiyenin %80'ini kullanmadığını gösterir.

Bu oranlar farklı çalışmalarda çalışmaların amacına göre kullanım kolaylığı sağlarlar. WOA ve WOE oranlarının toplamı 1 eder ve hep pozitiftirler.

İkinci tavsiye kullanım metodu olan regresyon tabanlı yöntemler ise klasik regresyon kullanarak tavsiyeyle son karar arasındaki ilişkiyi anlamaya yardımcı olur (örneğin , Harvey v.d., 2000; Hedlund v.d., 1998). Modelde ilk karar ve tavsiyenin son karar üzerindeki ağırlıkları regresyon betaları tarafından anlaşılır. Darlington (1968) ise tavsiye kullanımını ölçmek için önce son kararları ilk kararlarla regresyona sokmuş ve çıkan R kare sayısını not etmiştir. Daha sonra tavsiyeyi ekleyerek tekrar R kare sayısına bakmıştır. R kare sayısındaki yüzde artışa tavsiye kullanım oranı demiştir.

Son metod olarak doğrudan değerlendirme yaklaşımından bahsedilebilir. Doğrudan değerlendirme, farklı özelliklerin veya yeteneklerin seviyesini ölçmek için yaygın olarak kullanılan bir yöntemdir. Bu yöntemde Likert tipi ölçek kullanır ve Dawes (2008)'e göre 5 ve 7 ölçekli testlerde daha sağlıklı sonuçlar elde ediliyor. Bu yönteme göre genelde 1 kesinlikle katılmıyorum yada aşırı olumsuz anlamına gelirken en yüksek değerde ise kesinlikle katılıyorum yada olumlu anlamına gelmektedir. Tavsiye kullanımı konusunda yargıçların ne kadar güvenle karar verdikleri ölçülebilir. Direk olarak tavsiyeden ne kadar yararlandığını açıklaması da istenebilir.

Önceki bölümde belirtildiği gibi insanların ek bir görüşü ne kadar kullandığını hesaplamak için bilim adamları çeşitli ölçüm tekniklerini kullanmıştır. Tavsiye kullanımını ölçmekteki amaç bu tavsiyenin kararlarda iyileşmeye yol açıp açmadığını ölçmektir. Ancak, bu tavsiye almayı reddetmeyi tercih eden bir kişi için ölçülemez çünkü zaten tavsiyeden yararlanmamıştır. Fakat tavsiyeden uzaklaşanlar da kararlarını değiştirdiği

için onların performansı ölçülebilir. Tavsiye doğruluğunu ölçmek YDS çalışmaları, çeşitli ölçütler kullanmışlardır. Bunlar MSE (Ortalama Hata Kare), RMSE (kök ortalama kare hatası), MAE (Ortalama Mutlak Hata), MdAE (Medyan Mutlak Hata), MAPE (Ortalama Mutlak Yüzde Hata), MdAPE (Medyan Mutlak Yüzde Hata), RMSPE (kök ortalama kare Yüzde Hata), RMdSPE (Kök Ortalama Kare Yüzde Hata). Bunlara ek olarak ortalama farklılıklarından doğan sorunları ortadan kaldırmak için Hoover (2006) tarafından Ortalama Mutlak Hata'nın Ortalaması formülü çıkmıştır.

Bu tezde, insanların tavsiyeyi ne kadar kullandıklarını ve kullandıkları veya kullanmadıkları tavsiyenin onların kararlarını ne kadar iyileştirdiğini incelenecektir. Bunu yaparken, sınırlı bir hata ile bir yazılım tarafından oluşturulan rassal tavsiye kullanmayı tercih ettim. Diğer bir deyişle, tavsiye sabit bir standart sapma ile gerçek tahmin merkezli bir normal dağılımdan rasgele olarak seçildi. Rastgele tavsiye kullanımının ardındaki ana motivasyon rassal tavsiyelerin gerçek bir danışmanın tavsiyeleri gibi oluşturulabilmesi ve danışmanların genelde gelen tavsiyelerin kimden geldiğini bilememesi sebebiyle birbirinden farksız olmasıdır. Araştırmacılar genelde kolaylığı sebebiyle rassal tavsiye kullanmışlardır. Fakat önemli bir nokta hata payının büyüklüğüne göre amatör veya uzman bir danışmanmış gibi şekilendirilebilmesidir.

Anket kısmında rastgele tavsiye oluşturulduktan sonra, insanlara çeşitli sorular sorularak ilk tahminler alınmış, daha sonra tavsiyeler sunularak son kararları istenmiştir. Data kısmına geldiğimizde eşsiz ve yeni bir yaklaşımı takip ederek 3 kısma ayrılmıştır. Kararlarından uzaklaşarak tavsiyeyi aşanlar, tavsiyeye yaklaşanlar ve tavsiyeden uzaklaşanlar olarak 3 kısma ayrılmıştır. Örnek vermek gerekirse ilk kararı 50, tavsiyesi 100 olan bir

kararda yargıç 110'u tercih ederse 1. gruba; 80'i tercih ederse 2. Gruba ve 40'ı tercih ederse 3. gruba girmiş oluyor. Bu sınıflandırmadan sonra formül tabanlı , regresyon tabanlı ve ölçek bazlı yöntemlere göre 3 grubun tavsiye kulanım dereceleri hesaplanmıştır. Daha sonra da bu 3 grubun karar performanslarındaki gelişime bakılarak tavsiye kullananların kararlarını iyileştirip iyileştirmediği incelenecektir.

Deney konusuna geldiğimizde her bir karar verici için her soruda birer adet tavsiye verilmiştir. Anketler internet üzerinden yapılacak şekilde hazırlanmıştır (ekler kısmında bulunabilir). Üniversitelerin lisansüstü eğitimine devam eden 66 adet denek ankete katılmıştır. Ankette 20 adet soruyu cevaplandırmışlardır. Sorular ise televizyon, radyo veya internet gibi 20. Yüzyılda gerçekleşmiş 20 adet buluşun bulunuş yılından oluşmaktadır. Tahminler ise 1900 ile 1999 arasında olmak zorundadır. Ankette toplamda 66 kişiden 1320 adet ilk tahmin alınmıştır. Aynı adet ilk tahmin, tavsiye, son tahmin, gerçek tahmin ve 1-7 arasındaki ölçekler alınmıştır. Daha sonra 3 gruba ayırmak için tercihler ayrılmıştır. 300 adet data eşitlikler yüzünden elenmiştir. Eşitlikler ise ilk kararın son karara eşit olması, ilk kararın tavsiyeye eşit olması yada son kararın tavsiyeye denk olması yüzünden gerçekleşmiştir.

Sonuçlar göstermiştir ki 83 gözlemde tavsiyeden uzaklaşılmış; 213 gözlemde tavsiye aşılmış ve 626 gözlemde tavsiyeye yaklaşılmıştır. Bonaccio ve Dalal (2006) da son grubun daha çok gözlemleneceğini söylemiştir. Formül tabanlı sonuçlara göre tavsiyeye yaklaşanlar %70 oranında tavsiyeden yararlanırken tavsiyeyi aşanlar %250; tavsiyeden kaçanlar %-180 tavsiyeyi kullandılar. Regresyon sonuçlarına göre bu sayılar %67, %375 ve %-74 olarak gerçekleşti. Ölçek bazlı sonuçlara göre %51, %80 ve %18 olmuştur. Sonuçlar birbirine

yakın çıkmıştır. Tavsiye kullanmayanların oranı negatif çıkarken aşırı kullananların oranı da %100den fazla çıkmıştır. Tavsiye kullanlarda da kendi kararını %30 oranında kullandıkları gözlemlenmiştir. Tavsiyeden kaçanlar ise tavsiyeyi negatif olarak kullanmış gibi sayılmıştır. Burada ölçek bazlı sonuçlar için tüm ölçeklerin ortalaması alınmıştır. Fakat insanların ne kadar tavsiyeden yararlandıklarını kendilerine sorarak direk sonuçlara ulaşma hedeflenmiştir. Formül bazlı yöntemde de tek tek oranlar hesaplanıp ortalama alınmıştır. Regresyon yöntemi sonuç zenginliği açısından daha zengin bir yöntemdir. Modelin önem derecesine bakarak ne kadar bilgi taşıdığı kolayca gözlemlenebilmektedir. Yine de bu 3 model birbirine paralel sonuçlar vermiştir.

Tavsiye kullanım sonuçları böyleyken karar performanslarına bakmak için ortalama mutlak hata ve ortalama mutlak hataların ortalaması ölçüleri kullanılmıştır. Bu ölçüler YDS çalışmalarında daha önce genişce kullanılmıştır. Ortalama mutlak hatalara bakıldığında en fazla gelişim tavsiyeye yaklaşanlar ve tavsiyeyi aşanlarda gerçekleşmiştir. Fakat tavsiyeden uzaklaşanların karar performansları daha kötüye gitmiştir. Burada uzman tavsiyesine uymanın ne kadar önemli olduğunu görüyoruz. Bunun yanında 2 adet benchmark model oluşturulmuş ve bu 3 grupla kıyas edilmiştir. Bunların ilki sabit tahmin veren bir karar verici diğeri ise sürekli ilk tahminiyle tavsiyenin ortalamasını alan bir modeldir. Sonuçlara göre tüm gruplar sabit modelden daha başarılıyken ortalama model uzaklaşanlardan daha iyi görünmektedir. Tavsiyeyi aşanlarla ortalama modelin neredeyse aynı performansı gösterdiği gözlemlenmiştir.

Sonuç olarak karar verme literatüründe bireylerin kararları için tavsiyeyi nasıl kullandıklarını ölçmek için farklı kriterler geliştirilmiştir. Tavsiye

değerlendirilmesi ve kullanımı, danışma tutarlılığı, tavsiye sayısı ve hakimin ve danışmanlar arasındaki etkileşimin varlığı gibi çeşitli faktörlere bağlıdır. Bu tür çalışmalarda bazı faktörler dikkate alınarak karar verici mekanizmalara anket yöntemiyle bazı kararlar alması sağlanır. Tavsiye verilerek son kararını vermesi beklenir. Tavsiyeyi ne kadar kullandığını ölçmek için üç ana yöntem (formül tabanlı,regresyona dayalı ve ölçeğe dayalı) kullanılır. Önerileri kullanım dereceleri tüm formül tabanlı, regresyon temelli ve doğrudan değerlendirme tekniklerinde aynı yönde sonuçlar elde edilmesi sonucunu çıkarmıştır. Bununla birlikte, regresyon modelleri daha detaylı ve bilgilendirici bir şekilde yargıç ve tavsiye arasındaki etkileşimi açıklamaktadır. Öte yandan, formül- bazlı yaklaşımlar bir defalık kararının kullanımını ölçmek için daha elverişlidir.

Her YDS çalışmasında olduğu gibi, deney tasarımı ve metodolojisi hakkında bazı sınırlamalar vardı. Bu çalışmada danışman güvenilirliği ana konu olmadığı için hakimler ayrıntılı şekilde danışmanların uzmanlık ve bilgilerinden haberdar değildi. Ancak, bu sınırlama gelecekteki araştırmalar için rastgele tavsiyenin temsiliyetini ölçen çalışmalarda her karar sırasında bir danışmanın karar performansı hakkında geri bildirim verilerek yargıçların danışmanlar hakkında fikir sahibi olmaları sağlanabilir. Başka bir sınırlama ise örnekleme büyüklüğünün kısıtlı olmasıdır. Bu çalışmanın datası üç ayrı duruma ayrıldığı için daha fazla denek kullanmak daha sağlıklı sonuçlar almaya yardımcı olabilir.

Bir tavsiye ile karşı karşıya olduğunda karar vericiler, ilk kararı ile tavsiye arasında gelgitler yaşayabilir ve hangisine yakın olacağı konusunda muhakeme edebilir. Bu konuyu araştırmak için farklı durumlar ortaya atılmış ve üç ayrı davranış saptanmıştır. Tavsiyeye yaklaşanlar, tavsiyeden

uzaklaşanlar ve kendi kararından uzaklaşıp tavsiyeyi aşanlar olarak 3 ayrı kategori verilmiştir. Bu şekilde problem yaratabilecek durumları genel durumlardan ayırmış oluruz. Bu Bonaccio ve Dalal (2006 )'nın da çağrısıdır ki bu tür problem çıkaran davranışların temeline inmek önemlidir. Diğer bir konu ise formüle dayalı yöntemler hakkında önemli bir konu Yaniv ( 2004b)'nin formülüdür. Harvey ve Fischer (1997) formülünün mutlak değerini alarak bazı hiç tavsiyeden yararlanmamış danışmanları sanki tavsiyelerden aşırı yararlanmış gibi göstererek sonuçların abartılı çıkmasına sebep olmaktadır. Bu yüzden Harvey ve Fischer'ın formülünü kullanmak daha mantıklıdır.

Üçüncü kullanım yöntemi olan doğrudan değerlendirmeyi bu çalışmada sadece tavsiyeden ne kadar faydalandığını ifade etmeleri için kullanıldı. Fakat sonuçlar gösterdi ki karar vericilerin kullandıklarından daha azını ifade ettikleri gözlemlenmiştir. Yani örneğin tavsiyeyi %80 oranında kullandığı halde 1 ve 7 arasından 3'ü seçerek aslında daha fazla kullandığı halde az kullandığını söylemektedir. Bu tür bir davranış literatürde Yaniv ve Kleinberger (2000)'in benci yaklaşım olarak ortaya attığı insanların kendi kararlarına daha ağırlık vermesi şeklinde açıklanabilir. Cevaplanması gereken soru ise insanlar aslında böyle bir davranış sergilerken acaba gerçekten daha az mı göstermeyi tercih ediyorlar yoksa tavsiyeden faydalanma derecelerini mi bilmiyorlar. İleriki çalışmalarda bu konunun üzerine gidilerek cevap aranmalıdır.

Tavsiye literatüründe temel amaç tavsiye kullanımı sonucunda insanların kararlarında daha yüksek doğruluk elde etmektir. Bu noktada, tavsiyeden ve kendi tahmininden ne kadar yararlandığı karar performansını geliştirmede önemli bir rol oynar. Karar performansını ölçmek için bu tezde

ortalama mutlak hata ve ortalama mutlak hatanın ortalamaya oranı kullanıldı. Sonuçlar göstermiştir ki insanlar tavsiyeye yaklaştığı zaman veya tavsiyeyi aştığı zaman karar performanslarını arttırırken tavsiyeden uzaklaştıkları zaman daha kötü performans göstermişlerdir. Burada tavsiyenin uzman tavsiyesi olarak verilmesi önemli rol oynamıştır. Rassal tavsiye uzman olmayan tavsiye şeklinde olsaydı sonuç ne olurdu diye ilerikli çalışmalarda değinilebilecek konular arasındadır.

İlginç bir sonuç ise tavsiyeden uzaklaşanların ilk karar performansı diğerlerinden daha iyidir. Bu durum o insanların tavsiyeden uzaklaşmalarına sebep olabilir mi diye düşünürken iyi performans göstermeleri onların kendine güvenini arttırmış olabilir ve bu yüzden tavsiyeden uzaklaşmış olabilirler. İleriki çalışmalarda bu tür insanların ilk kararlarında duydukları güveni ölçmek daha faydalı olacaktır. Tersten baktığımızda da tavsiyeyi aşanların ilk performansları daha kötüdür. Bu yüzden tavsiyeyi fazlasıyla kulandıklarını gözlemlemekteyiz. Burada araştırılması gereken konu, özgüvenle tavsiyeden yararlanma arasındaki ilişki ve ilk performansla özgüven arasındaki ilişkidir. Kısacası, tavsiye alma literatüründe analiz edilmesi gereken bazı değinilmemiş konular vardır. İleriki çalışmalarda bu konulara değinmek birçok sorunun cevabını bulmaya yardımcı olacaktır.

# APPENDIX C: TEZ FOTOKOPİSİ İZİN FORMU

**ENSTİTÜ**

Fen Bilimleri Enstitüsü ☐

Sosyal Bilimler Enstitüsü ☑

Uygulamalı Matematik Enstitüsü ☐

Enformatik Enstitüsü ☐

Deniz Bilimleri Enstitüsü ☐

**YAZARIN**

Soyadı : ÖZARSLAN
Adı : Ali
Bölümü : İşletme

**TEZİN ADI**: "Advice Utilization and Decision Performance under Random Advice"

**TEZİN TÜRÜ** : Yüksek Lisans ☑ Doktora ☐

1. Tezimin tamamından kaynak gösterilmek şartıyla fotokopi alınabilir. ☑

2. Tezimin içindekiler sayfası, özet, indeks sayfalarından ve/veya bir bölümünden kaynak gösterilmek şartıyla fotokopi alınabilir. ☐

3. Tezimden bir bir (1) yıl süreyle fotokopi alınamaz. ☐

**TEZİN KÜTÜPHANEYE TESLİM TARİHİ**: