8-2010

Who is Worth What? Judge and Advisor Characteristics in a Paid-Advice Judgment Scenario

Benjamin Slade
Clemson University, bslade@clemson.edu

WHO IS WORTH WHAT?
JUDGE AND ADVISOR CHARACTERISTICS
IN A PAID-ADVICE
JUDGMENT SCENARIO

A Thesis
Presented to
the Graduate School
of Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Applied Psychology

---

by
Benjamin H. Slade
August 2010

Accepted by:
Dr. Patrick Raymark, Committee Chair
Dr. Christopher C. Pagano
Dr. Fred S. Switzer, III

ABSTRACT

When making a decision, people often receive advice before settling on a particular course of action. Decision makers exhibit a spectrum of responses to advice, ranging from total rejection to complete acceptance. The purpose of this study is to examine predictors of advice use within a Judge-Advisor System (JAS: Sniezek & Buckley, 1995). Prior research has examined a variety of task characteristics, advisor characteristics, and decision maker characteristics (e.g. Bonaccio, 2007; Gino & Moore, 2007; Yaniv & Kleinberger, 2000). In this study, judge characteristics including confidence, accuracy, prior task knowledge, and other individual differences are examined. Advisor characteristics such as advice cost and advisor expertise are manipulated between persons. Results indicated that judge accuracy, confidence, and prior task knowledge were all negatively related to advice utilization. Advice was weighted more heavily when the judge was told that the advice came from an expert than from a novice, but advice cost did not influence advice use. Implications and future research directions are discussed.

DEDICATION

This thesis is dedicated to Dr. Allen Slade. He has been a source of wisdom for countless practical, technical, and spiritual matters. As a kid, I never expected to follow so closely in your footsteps, but ‘talking shop’ with you has been incredibly rewarding throughout my graduate school journey. Thanks, Dad.

TABLE OF CONTENTS Page

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Real decision problems rarely contain all the information that is necessary to solve them. People therefore engage in social sense-making processes in order to understand situational cues (Yates, Price, Lee, & Ramirez, 1996; Zarnoth & Sniezek, 1997). Decision makers ("Judges") often solicit the opinion of others ("Advisors") before making a final commitment. A politician may receive advice from environmental scientists before voting on legislation about sustainable resources, a high school student queries peers and parents prior to choosing a college, and a lost motorist may request advice from a pedestrian about the best way to reach a given destination.

Although advice is an important part of decision making, advice research is a relatively new field (Bonaccio & Dalal, 2006). Most decision research has neglected the social context of decisions (Payne, Bettman, & Johnson, 1993). The research on 'small groups' decision making has primarily focused on persons with undifferentiated roles, even though many groups have a primary decision maker (Kerr & Tinsdale, 2004). There are relatively few empirical or theoretical studies linking advice to decision making (e.g. Dalal, 2001; Patt, Bowles & Cash, 2006; Schrah, 2000; Sniezek, Schrah, & Dalal, 2004; Yaniv, 2004; Yaniv & Kleinberger, 2000).

When a judge seeks advice from a qualified advisor or advisors, he or she assesses the quality of the advice and integrates the advice with his own evaluation to make a final decision (Yaniv, 2004). Receiving advice generally improves decision accuracy (Patt, et al., 2006; Sniezek, et al., 2004; Yaniv, 2004; Yaniv & Kleinberger, 2000). Seeking advice can help the judge fill in missing information, consider alternative options, or confirm an existing opinion. Advice use improves decision quality when advisors are ‘experts’ that have received task specific training above and beyond the judges (Harvey & Fischer, 1995; Schrah, 2000). Moreover, receiving advice improves the accuracy of decisions even when the advisor and judge are equally qualified on the subject (Gino, 2008; Gino & Moore, 2007; Yaniv & Kleinberger, 2000). When a judge averages more than one independent estimate into a final decision, it tends to reduce error, even if the advisor is no more accurate than the judge (Larrick & Soll, 2006).

Even though advisors help people make better decisions, judges are not always very good at knowing when and how much to use the advice. This paper will explore several factors which influence the extent to which people use advice. After giving a general overview of Judge-Advisor System (JAS) scenarios, I will define egocentric bias and explain how it can paint an inordinately rosy picture of the judge’s own knowledge and accuracy (Yaniv & Kleinberger, 2000). Furthermore, I will discuss individual differences that may influence a judge’s receptiveness to advice, including decision-specific knowledge, agreeableness, and confidence (Bonaccio, 2007; Bonaccio & Schrah, 2006), as well as the relationship between judge confidence and accuracy (Yaniv, Choshen-Hillel, & Milyavsky, 2009).

After exploring judge characteristics, the paper will address characteristics of the advice and advisor. Judges adjust their advice use accordingly to their perceptions of the advisor and the perceived quality of the advice (Harvey & Fischer, 1995; Yaniv & Kleinberger, 2000). Finally, prior research has found that advice cost strongly influences advice use (e.g. Gino, 2008). However, the mechanisms behind this paid advice effect are incompletely understood. As sunk costs theory is the prevailing explanation of the paid advice effect, it will be addressed in some detail, along with several competing and complementary theories (Arkes & Blumer, 1985; Gino, 2007; Monroe, 1973).

Judge-Advisor Systems

Advice studies often use a Judge-Advisor System (JAS) paradigm to create pertinent decision-making scenarios. The participant (or judge) is responsible for giving the best possible answer to a decision problem. After making an initial judgment, the judge receives advice from one or more advisors. The judge is then responsible for integrating the judgments of the advisors with his own initial judgment, and then making a final decision. This paradigm accurately mimics many real-world advice scenarios. Furthermore, because it collects judgment information both before and after receiving advice, it is possible to assess the extent to which a judge modifies his or her final decision after receiving advice.

There are two primary JAS paradigms. In ‘choice’ scenarios, judges are presented with multiple options and asked to select the best one. Researchers have generally used product decisions for choice scenarios, such as identifying the most valuable mountain bike (Schrah, 2000) or hiking backpack (Schrah, Dalal, & Sniezek, 2006). These qualitative scenarios mimic decision making tasks such as choosing between medical treatment options or purchasing a product. Advice in these scenarios generally consists of a recommendation for a specific option. ‘Judgment’ scenarios, by contrast, consist of one or more quantitative estimations. Advice taking studies have used quantitative estimates of such divergent topics as historical dates (e.g. Gino, 2008; Yaniv & Kleinberger, 2000), sales forecasts (Harvey, Harries, & Fischer, 2000), a person's body weight based on a photograph (Gino & Moore, 2007), and livestock mortality rates for different epidemics (Harvey & Fischer, 1997). After a judge has made an initial estimation, the advisor's estimate is made available and the judge is asked to make a final decision. One advantage that judgment scenarios offer is the ability of judges to accept advice in degree rather than kind. Judges can be influenced by advice without fully accepting or fully rejecting it. The weight of advice (WOA) can then be calculated to evaluate advice utilization. WOA for a final estimate ranges from zero (when the final estimate is exactly the same as the initial estimate) to one (when the final estimate is exactly the same as the advisor's estimate). This study will use a quantitative judgment scenario in order to evaluate how judges revise their opinions based on the advisor's recommendation.

Egocentric Bias

Judges vary their weighting policies based on their perceptions of themselves and their perceptions of the advice[1]. Judges will lend more weight to an advisor's recommendation if they are not confident in their own answers (Gino, 2008). Also, a judge forms an opinion about the quality of the advice based on the perceived motivation and expertise of the advisor, and varies his weighting policy accordingly (Bonaccio & Dalal, 2009). This opinion can change across multiple interactions as a trust relationship is developed (Yaniv & Kleinberger, 2000).

[1] In the current context, a weighting policy is defined as the extent to which a judge uses an advisor's estimate as compared to her own estimate when making a final decision.

Individuals rarely use an optimal weighting strategy. Instead, all other things being equal, they give disproportional weight to their own opinions as compared to outside opinions. This non-optimal weighting error is known as advice discounting. Yaniv and Kleinberger (2000) examined advice discounting in several judge-advisor studies (JAS), using estimates of the dates of various historical events as the dependent variable. They found that while judges are indeed more accurate when they have access to advice, they do not use an optimal weighting strategy, discounting the advice of others relative to their own opinions. Even though the initial estimates for judges and advisors had similar accuracy ratings, the mean observed weight of own advice (WOA) was .29, indicating non-optimal discounting.[2]

There are several reasons why an individual may underweight outside advice. If a judge perceives that he is better informed or a more qualified decision maker than the advisor, he will therefore give additional weight to his own judgments (Kreuger, 2003). If a judge perceives himself as less qualified than the advisor, he will give less weight to his own judgments. If everyone had an accurate perception of their own abilities, and were randomly given access to the estimate of another person, it should result in an average WOA of 0.5 (Larrick & Soll, 2006). However, people generally consider their own opinions and judgments to be more accurate than the 'average' person. This egocentric bias leads to people generally overestimating the accuracy of their own opinions, and skews weighting policies in a non-optimal manner (Kreuger, 2003).

[2] While most research studies use Weight Of Advice (WOA) to describe advice utilization, some studies use Weight of Own Estimate (WOE). WOE, while also ranging from 0 to 1, is the opposite of weight of advice (e.g. WOE for this study was .71). All future WOE will all be converted to WOA.

Economic theory supports the egocentric bias explanation of advice discounting (e.g. Nishimura, 1991). Whenever a firm or individual has more information about their own capital then one does about a different party, there is an imbalance in knowledge known as differential information. Whereas a corporation may have intimate knowledge about its own investments, technology, and strategy, they will rarely have such insight into the designs of other corporations. Compounding this problem, people generally do not 'assume the best' when it comes to missing information. For example, judges tended to interpret missing information about an advisor's intentions or expertise negatively (Bonaccio, 2007). When applied to judge-advisor scenarios, this information imbalance may encourage judges underweight advice. If they do not have access to the 'private knowledge' of the advisor, they will place more trust in their own opinions (Yaniv & Kleinberger, 2000).

Decision makers are privy to their own reasons for holding an opinion or making a choice, but do not have the same access to an advisor's internal reasoning process (Yaniv & Kleinberger, 2000). Since judges have less information about the logic or motivation of the advisor as compared to the self, they discount the advice accordingly. This problem may be attenuated in JAS studies, which generally convey the advisor's opinion in a sparse fashion (e.g. a text box on a computer screen stating "The advisor's estimate is X"). It is possible that the differential information problem is accentuated by using a relatively information-poor way of receiving advice in such decision-making studies. In the real world, advisors are more likely to explain their rationale for making a particular recommendation or holding an opinion. It is possible that this egocentric bias is not as substantial outside of the laboratory.

Anchoring theory (Tversky & Kahneman, 1986) may also shed some light on advice discounting. In most advice studies, judges give an initial estimate before receiving input from an advisor. This initial estimate then serves as an anchor from which the judge then adjusts his final estimate. By asking judges to make a formal estimate prior to receiving advice, it is possible that it solidifies the judges' commitment to their initial estimate. On average, judges are unwilling to adjust their scores too far away from their anchor, which thus leads to advice discounting when the advisor's estimate is far away from the judge's initial estimate (Harvey & Fischer, 1997).

Individual differences in egocentric discounting

However, people do not always favor their own opinions when faced with a decision task. For example, egocentric discounting is reduced when there are financial incentives to do well, or when making a decision that could have negative consequences (Bonaccio & Dalal, 2009; Dalal, 2001; Sniezek et al, 2004). By using advice, judges avoid shouldering the full responsibility for the final decision. Participants vary their weighting strategies according to the difficulty of the task, even when advisor quality stays constant (Gino & Moore, 2007). As noted earlier, when an advisor is roughly as knowledgeable as the judge, the optimal strategy is to adapt a weighting policy of .50. However, people tend to overvalue their own opinion when the decision task is relatively simple, and overvalue the opinion of their advisor when a decision task is relatively difficult, even if the advisor is equally qualified to give advice on both the difficult and the simple tasks (Gino & Moore, 2007). Judges are also less likely to spend their own time acquiring information for a decision task when the task is difficult and they can consult an advisor (Payne, Bettman, & Johnson, 1988). Task difficulty therefore moderates the relationship between an advisor's recommendation and the judge's final estimate.

Judges also make assessments of their own expertise (Arkes, Christensen, Lai, & Blumer, 1987; Trafimow & Sniezek, 1994). Arkes et al. (1987) were able to experimentally manipulate the participants' perception of content-specific expertise. Accordingly, participants that had their perceptions of expertise bolstered reported more confidence answering a set of questions, whereas participants in the low-expertise manipulation reported less confidence while answering the same set of questions (Arkes et al, 1987; Trafimow & Sniezek, 1994). Judge confidence in their own estimate is a significant predictor of advice use (Gino, 2008). If a judge does not believe his answer is accurate, he will be more receptive to the recommendation of an advisor. Additionally, there is a positive relationship between a judge's initial estimate error and weighting policy, such that judges with inaccurate initial judgments are more likely to utilize an advisor's opinion (Yaniv & Kleinberger, 2000).

There has been limited research on the role individual differences play on judge-advisor relationships. One individual difference which is theoretically linked to advice utilization is the agreeableness facet of the Five-Factor Model (FFM; Mount & Barrick, 1995). Agreeable individuals are generally described as friendly, supportive, and altruistic people that are likely to follow the desires and wishes of others. A literature review did not turn up any peer-reviewed sources specifically linking agreeableness to advice use. However, prior research has linked judge expressivity to advice utilization (Feng & MacGeorge, 2006). Judges that are more expressive are more likely to use advice. Expressivity is defined in part by emotionality and interdependence, and is moderately correlated with agreeableness (Gross & John, 1995). Additionally, agreeableness is positively related to consensus decision making and supportive communication (Sager & Gastil, 2006), suggesting that highly agreeable decision makers are more likely to incorporate the opinions of others. Moreover, Bonaccio's (2007) dissertation found agreeableness to be a significant predictor of advice utilization. Specifically, Bonaccio found an interaction between agreeableness and advisor expertise on advice use, such that high-agreeableness individuals were more likely to incorporate estimates from novice advisors into their final estimate. Generalized self-efficacy has a positive relationship with performance as well (Bandura, 1982) although some studies suggest that the relationship between self efficacy and performance is not positive in all situations (e.g. Vancouver, Thompson, Tischner, & Putka, 2002).

Research findings are mixed on the link between decision confidence and decision accuracy. Some studies suggest that there is a positive relationship between confidence and decision accuracy. For example, confident eyewitnesses are generally more accurate (Sauerland & Sporer, 2009). Confidence and performance both tend to increase with experience (Harvey & Fischer, 1995). However, confidence is not always a valid predictor of accuracy (Lichtenstein, Fischoff, & Phillips, 1982). In one study, some judges were given estimates randomly generated from a pool of responses, whereas other judges were given the estimates from the pool of responses that were most similar to their own (Yaniv, Choshen-Hillel, & Milyavsky, 2009). Even though judges were told how their advice was generated, judges were more confident in their final estimates when told they had received the estimates which were closest to their own opinions (interdependent), even though this gives them little meaningful information. When given randomly sampled (independent) estimates, their confidence decreased, even though this new information lead to more accurate estimates (Yaniv et al, 2009). Independent estimates are more useful for improving decision accuracy (Yaniv, 2004), but people report more confidence in their final estimate after receiving advice which is similar to their initial estimate (Yaniv et al, 2009).

Harvey, Harries, and Fischer (2000) found that people are better at assessing the quality of an advisor's estimate than actually utilizing it. Judges were asked to forecast monthly sales, and then given forecasts from four other advisors. After assessing the accuracy of each advisor's forecast, the judge then had to incorporate the advice into a final estimate. Individuals were better at estimating the quality of the advisor's advice then they were at compiling all estimates into a final sales forecast. One possible explanation for this is that integrating the advice puts a heavier load on the judge's working memory. Even though participants were able to accurately assess the quality of the advice, they may have experienced some difficulty holding all forecasts in working memory while applying a weighting policy, and therefore used non-optimal weighting strategies. This effect decreases when the advisors are distinguished more saliently, perhaps easing the load on the judge's working memory (Harvey et al, 2000).

Judges are more willing to pay a fee for good advice than poor advice, and are able to make an accurate determination of advice quality in as little as three trials (Yaniv & Kleinberger, 2000). There is also less egocentric discounting when the advisors are knowledgeable experts (Sniezek & Buckley, 1995). Feng and MacGeorge found that judges place more value on advice when the advisor is older, more experienced, more educated, and wiser than themselves (2006). Judges also are more likely to listen to advice from people with whom they already have a close relationship (Feng & MacGeorge, 2006). Judges prefer advice from advisors that exhibit more confidence in their answers and that have a better track record of success (Sniezek & Buckley, 1995). Confidence may act as a signal of quality because both confidence and performance tend to increase with experience (Harvey & Fischer, 1997).

Judges lend more weight to good advice than poor advice; Yaniv and Kleinberger (2000) found WOA averages of .52 and .26, respectively for good advice and poor advice. However, even when advice was good, it seems the participants were still not using it as much as they should. In the good advice condition, the advisor was more accurate than the judge's initial judgment in 97% of the cases, so the judge's optimal WOA should have been much greater than .50.

Research has shown that people prefer expert advisors to novices (Sniezek & Buckley, 1995). However, to my knowledge, only one research study has manipulated advisor characteristics and 'expertise' while holding the actual advice quality constant (Harvey & Fischer, 1997).

Paying for Advice

Another variable which influences egocentric discounting is the cost of advice. Most advice studies have given participants advice for free (e.g., Harvey & Fischer, 1995; Yaniv, 2004; Yaniv & Kleinberger, 2000). However, several studies have been conducted where participants pay to receive advice. These studies have found that people treat advice as more valuable and more accurate if they had to pay to receive it (Gino, 2008; Patt, et al. 2006; Sniezak, et al. 2004). Because the judge has paid for the advice, he or she perceives it as valuable and thus takes it into greater account while making a final estimate. This over-valuation of paid advice relative to free advice of the same quality is known as the paid-advice effect. The paid-advice effect has far-reaching implications because people generally prepay for advice long before they are able to accurately assess its quality, and before they choose how to utilize the advice they have been given (e.g. financial advisors, management consultants, marriage counselors).

People weigh paid advice more heavily than free advice even when they are informed that the costly advice is of the same quality as the free advice (Gino, 2008). Participants were twice given the option of receiving advice in a JAS scenario, estimating U.S. historical dates. They were offered a block of free advice for one set of questions and could give up a certain amount of their winnings for advice on the other block of questions. The participants were told that advice in both blocks consisted of estimations from randomly selected participants from the pilot study. This meant that advisors were, on average, as equally qualified as participants for the judgment task. Moreover, because advisors were randomly selected for each question, it was impossible for judges to form an impression of advice quality across questions. Gino (2008) found that the mean WOA was higher when participants paid for advice than when they received it for free (WOA=.72 and .46, respectively). Participants actually overweighted the advice of the advisor when they had to pay for it, and slightly underweighted it when they received it for free.

Sniezak, et al. (2004) also found that prepaying for advice makes participants more likely to utilize the advice in their final decision. Moreover, because the egocentric bias was reduced, the participants in the paid advice condition were actually more accurate with their final estimates. Because receiving the advice had cost them something, they were less likely to reject the advice in favor of their own estimation. A follow-up study suggested that the relationship between advice prepayment and advice use is mediated by perceptions of advisor credibility (Patt, et al. 2006). Judges trusted the advisor more when they had prepaid for the advice, and therefore were more likely to accept the advisor's judgments as accurate.

Since prepaying for advice increases the perceived credibility of an advisor, it is important to examine why people are willing to pay for the advice. It is possible that paying for advice may inappropriately increase advice use and judge confidence in an advisor. Although the paid advice effect reduces the egocentric bias, decision-makers may simply be swapping one decision error for another. There are many circumstances when people are willing to pay for the opinion of others, even if there is a shortage of empirical evidence that the advice is worthwhile. Professional investors charge high fees for their services, even though the stock market is inherently unpredictable (Bogle, 1999). Likewise, people purchase mutual fund recommendations from fund managers, even though the fund managers have an inherent conflict of interest in their recommendations (Freeman & Brown, 2001). Businesses hire external consultants to solve management problems, and individuals pay for therapists to solve personal problems, although neither one appears to be all that helpful (Dawes, 1994; Micklethwait & Woolridge, 1996; Pfeffer & Sutton, 2006).

Judges are more receptive to the opinions of an advisor (and therefore more willing to pay for said opinions) when a task is complex versus a simple task, even if the advisor is no more qualified than the judge (Gino & Moore, 2007). People are generally more willing to pay for advice when they are less confident in their own responses (Godek & Murray, 2008). Research has shown that people generally believe they are above-average on tasks which most people are able to do (such as driving a car or operating a computer mouse), but below average on more challenging tasks, such as computer programming (Kruger, 1999). Gino and Moore found that people are over-confident in their decision accuracy on simple tasks and under-confident in their decision accuracy on complex 'judgment' tasks (Gino & Moore, 2007). Although some research suggests the opposite is true for 'choice' tasks (Lichtenstein & Fischoff, 1977), this could be an artifact of the experimental design.

Decision specific knowledge is another important predictor of willingness to pay for advice. When they are low on decision-specific knowledge, judges are more willing to solicit opinions and additional information, even if it comes at a cost to themselves (Godek & Murray, 2008). Decision makers that use rational processing approaches generally are willing to pay more for advice than individuals that are making decisions based on affective and emotional intuitions, also known as experiential processing (Godek & Murray, 2008). However, this relationship is moderated by decision specific knowledge, such that rational processors are only willing to pay more for advice when they perceive their own knowledge to be minimal. When the judge has a large amount of decision specific knowledge, there is no significant difference between rational processing and experiential processing. Judges that know the subject area are less willing to pay for outside opinions, regardless of their cognitive processing strategy.

One possible explanation for the overuse of prepaid advice is that judges who pay for such advice are incapable or unwilling to rely on their own estimate in a final decision. However, it seems that the paid-advice effect is more than just an indicator of (a lack of) self-confidence or an idiosyncrasy of individuals that are willing to pay for advice. In Study 2 of Gino (2008), all participants were required to receive advice in both the paid block of advice and the free block, eliminating any self-selection effects for advice payment. The effect was just as robust even when they had no choice whether or not they paid for the advice. This suggests that the impact advice cost has on advice use is causal and psychological in nature, not just a signal of confidence. Secondly, participants were asked to give 90% confidence intervals of their initial estimates. The range of the confidence interval was a significant predictor of advice use such that judges were more likely to utilize an advisor's judgment on questions where the judge reported a wide confidence interval. In other words, when the judge was merely guessing the right answer, she was eager to use the advice she received. However, the effect of advice cost on advice use was significant even when controlling for judge confidence. Judges gave more weight to prepaid advice then they should have, given the confidence expressed in their initial estimate (Gino, 2008). Paid advice was valued more than unpaid advice even when the judge was confident in her answer.

Sunk Costs

The prevailing view in advice research is that the paid advice effect is an extension of the sunk costs effect (Gino, 2008). Because paying for advice is an investment decision that yields a future product (namely, purchasing advice for one or more decision-making problems), it is important to discuss investment research and sunk costs in some detail. Rational economic decisions should not include prior investments, but only evaluate the relative costs and benefits of future behavior. However, people do not generally behave rationally in their investment decisions. Individuals are more likely to continue spending resources on a project if they have already invested significant time, money or effort into the endeavor (Arkes & Blumer, 1985). This holds true even when the project is futile. This maladaptive tendency to throw good money after bad is known as the sunk costs effect.

Garland found a linear relationship between sunk costs on a research and development (R & D) project and commitment to complete it (1990). Participants that had already poured a sizable investment into a project chose to spend additional money to complete it even after a rival company came out with technology that rendered the project obsolete (Garland, 1990). Decision makers may try to justify their past expenses by completing projects, even if the costs to complete the project outweigh the benefits (Staw, 1981). This attempt at justification can be targeted towards others who might evaluate their performance (e.g. stakeholders, supervisors, coworkers, the general public). Additionally, the prevalence of the sunk-cost effect in personal decisions (ranging from auto repair to romantic relationships) suggests that these attempts at justification are self-targeted as well.

This irrational economic behavior may extend from an overgeneralization of the "Don't waste" rule (Arkes & Ayton, 1999). Since people see themselves as fundamentally rational beings that only expend resources to gain something, they may be willing to expend a large amount of resources to complete an ongoing project that is almost certainly doomed, rather than accept a loss with no return. Arkes (1996) asked participants to respond to a vignette about an R & D company that is developing a new tent polymer. When a rival company began marketing a far superior product, the decision-maker had to choose between continuing the investment and abandoning the sunk cost. The researcher found that people were more willing to abort the failing project if they sold the leftover material to a roofer for $5000 then if they sold it as 'scrap,' also for $5000. Since people do not want to appear wasteful, and associate scrap with waste, they chose to continue down an ill-advised investment path rather than cut their losses.

The sunk costs effect influences future economic investment based on prior economic investment. It also influences future behavior based on prior economic investment. Arkes and Blumer (1985, Experiment 2) experimentally manipulated the cost of season tickets for the Ohio University Theater. Patrons who paid the full price for their season tickets were more likely to attend performances, whereas patrons that received a discount attended fewer events. Once a season pass had been purchased, all ticket-holders had an identical license to attend any and all plays hosted by the theater. Because the cost of tickets was varied randomly across participants, one can assume that the costs and benefits of attending a show would be roughly equal across groups. The only difference was the amount of money each person had 'sunk' into that license. Therefore, the size of the prior investment played a significant role in determining whether patrons would attend or not attend each show. Likewise, in a situation where a judge has prepaid for advice, if he has already 'sunk' money into that advice, the size of that investment may play a role in the way he uses that advice in the future (Gino, 2008).

Likewise, individuals acted irrationally when given the following scenario:

> *Assume that you spent $100 on a ticket for a weekend ski trip to Michigan. Several weeks later you buy a $50 ticket for a weekend ski trip to Wisconsin. You think you will enjoy the Wisconsin ski trip more than the Michigan ski trip. As you are putting your just-purchased Wisconsin ski trip ticket in your wallet you notice that the Michigan ski trip and the Wisconsin ski trip are for the same weekend. It's too late to sell either ticket, and you cannot return either one. You must use one ticket and not the other. Which ski trip will you go on? (Arkes & Blumer, 1985, p. 126).*

Even though the scenario stated that the decision maker expected to enjoy the Wisconsin ski trip more than the Michigan ski trip, over half the participants chose to attend the more expensive Michigan vacation. Because they had already paid for both trips, the cost of attending should no longer be a consideration in choosing which trip. However, most participants chose the option that had cost more in the past but would be less enjoyable in the future. Over generalizing the rule to avoid or minimize waste erroneously leads the participants to make an illogical choice. In a judge-advisor scenario, the judge may try to minimize waste by using advice more if she has paid for it.

If the advice is free, it is reasonable to reject advice as being poor quality (or average quality) without invoking an impression of waste, but if advice is costly, rejecting advice is wasteful. Moreover, the strength of the ‘don’t waste’ rule increases as the cost of the advice increases (Gino, 2008).

Although the sunk costs effect is the dominant explanation for the paid advice effect, there are two complementary explanations which may supplement sunk costs theory in explaining why paying for advice causes individuals to weight the advice more heavily. Judges may use advice cost as a signal of advice quality and therefore adjust their weighting policy accordingly. In the absence of other information, consumers will assume that more expensive products are more valuable than inexpensive products (Monroe, 1973). Although the correlation between price and quality varies across product categories (Ordonez, 1998), the relationship is generally positive. Patt, et al. (2006) found that judges that prepaid for advice considered their advisors to be a more credible source of advice. Accordingly, judges may therefore associate advice cost with advice quality, and therefore weigh paid advice more heavily.

Cognitive dissonance may also play a role in this perception of advisor credibility (Festinger, 1957; Gino, 2008). People experience cognitive dissonance if they pay for something but do not use it. Judges generally perceive themselves as someone who only gives money for something which has value. Since judges generally believe themselves to be rational and see themselves as one who would only pay for advice if it is valuable, they may experience discomfort if they pay for advice but choose to ultimately reject the advice. If they have paid for advice, they must either use the advice in order to demonstrate its value, or change their schema of themselves to recognize that they have paid for something worthless. In order to relieve this dissonance, judges that have paid for advice may change their weighting policy in order to treat the advice as valuable.

To my knowledge, Gino (Study 3: 2008) is the only study which has directly examined the psychological mechanisms behind the paid advice effect. This study experimentally manipulated sunk costs, such that participants either had to pay $1 or $2 for advice. Half the participants in the $1 condition ended up receiving the advice for free, as did half the participants in the $2 condition. The study supported the sunk costs effect as a mechanism for the paid advice effect. Gino (2008) did not find judge perception of advice quality to be a significant predictor of advice use. However, she was using a study design that explicitly told judges that the advice was of the same quality regardless of how much one paid for the advice. This may have confounded her measure of advice quality. Additionally, Gino (2008) did not find self-reported cognitive dissonance to be significantly related to advice use. However, because cognitive dissonance is a subconscious process, it is possible that the participants were not aware of any cognitive dissonance, or would have experienced even more discomfort in reporting this dissonance. If a judge has indeed changed their schema or weighting policy in order to see the advice as valuable, he would be unlikely to still be experiencing cognitive dissonance and thus would be unlikely to report it.

Hypotheses

Judges do not use advice optimally (Yaniv & Kleinberger, 2000). However, almost any decision strategy which incorporates more than one independent observation will yield more accurate answers, because averaging cancels out error (Larrick & Soll, 2006). Therefore, even though judges may use imperfect weighting policies, it can be expected that participants will generally estimate more accurately after they have received advice.

Hypothesis 1: Final estimates will be more accurate than initial estimates

Judges are able to make an evaluation of accuracy, even without feedback (Yaniv & Kleinberger, 2000). Although self report is not always the most reliable indicator of competency, it does correlate with objective measures of performance (Jonas & Fletcher, 2004). Regardless of the accuracy of their self-evaluations, judges that consider themselves to be 'experts' should be less likely to use advice, and judges that consider themselves to be relatively ignorant of the subject will be more receptive to advice. This has been examined at an item level before (e.g. Gino, 2008), but to my knowledge there has been no empirical research on global competency ratings and advice use.

Hypothesis 2a: Self-ratings of U.S. History knowledge will be positively related to accuracy.

Hypothesis 2b: Self-ratings of U.S. History knowledge will be negatively related to advice use.

Agreeable individuals are more likely to seek consensus decisions and compromise with others (Sager & Gastil, 2006). Accordingly, they will be more willing to shift their own estimates to take into account the opinion of an advisor. Bonaccio (2007) found agreeableness to predict advice utilization, but this has not yet been replicated in the advice literature.

Hypothesis 3: Agreeableness will be positively related to advice use.

Judges vary their WOA based on the perceived quality of the advice (Yaniv & Kleinberger, 2000). Research has shown that people prefer expert advisors to novices, and expert advice reduces egocentric discounting (Sniezek & Buckley). Judges prefer advice from more experience, more educated, and wiser advisors (Feng & MacGeorge, 2006). However, no research has been done which manipulates advisor 'expertise' while holding the actual advice quality constant.

Hypothesis 4a: Participants will use 'expert' advice more than 'novice' advice, even when holding advice quality constant.

Hypothesis 4b: Participants will perceive 'expert' advice as more valuable than 'novice' advice, even when holding advice quality constant.

Hypothesis 4c: Participants will report greater confidence in their final estimates when they are receiving expert advice then when they are receiving novice advice.

Judges that pay an advisor will weight that advice more heavily than judges that receive free advice (Gino 2008). The prevailing theory is that sunk costs explain this increase in advice utilization. However, the cost of advice may also act as signal of the advice quality (Monroe, 1973). If advice cost does indeed act as a signal of quality to the judge, it is reasonable to expect that advice which has a price tag attached will have a high WOA and be rated as more valuable. Moreover, this will be true even if the judge is told he or she does not have to pay for the advice.

Hypothesis 5a: Judges that pay for advice will use it more than judges that receive advice for free.

Hypothesis 5b: There will be a significant difference in advice use between the 'free' condition and the 'costly no payment' condition, such that the 'costly, no payment' advice is used more than the 'free' advice. [3]

Hypothesis 5c: Participants will report a higher perceived value for the 'costly, no payment' condition than for the 'free' condition.

People generally overestimate their own competency, especially on simple tasks (Gino & Moore, 2005). Participants that are very confident in their own responses are unlikely to be swayed by an advisor's recommendation, and may only be looking for confirmation of the decision they reached during their 'hypothesis generation' phase (Schrah, Sniezek, & Dalal, 2006). Likewise, participants that have little confidence in their initial estimate can be expected to weigh the advisor's recommendation more heavily (Gino, 2008).

Accordingly, participants that had their perceptions of expertise bolstered reported more confidence answering a set of questions, whereas participants in the low-expertise manipulation reported less confidence while answering the same set of questions (Arkes et al, 1987; Trafimow & Sniezek, 1994). Judge confidence in their own estimate is a significant predictor of advice use (Gino, 2008). If a judge does not believe his answer is accurate, he will be more receptive to the recommendation of an advisor. Additionally, there is a positive relationship between a judge's initial estimate error and weighting policy, such that judges with inaccurate initial judgments are more likely to utilize an advisor's opinion (Yaniv & Kleinberger, 2000).

[3] In the 'costly no payment' condition, the participant is informed that the advice costs a certain amount, but they do not have to pay for it. In the 'free' condition, the cost of the advice is not mentioned. See the procedure section for a more complete explanation.

Hypothesis 6a: Judge confidence in his/her initial estimate will be positively related to accuracy.

Hypothesis 6b: Judge confidence in his/her initial estimate will be negatively related to advice use.

Hypothesis 6c: There will be a stronger negative relationship between judge confidence and advice use in the novice advisor condition than in the expert advisor condition.

Judges are generally more receptive to information which supports their position rather than contradicting it, even if such information yields little new information (Yaniv & Milyavsky, 2009). Judges who receive advice that is similar to their own estimate are therefore more likely to regard their estimate (and the advice) as accurate.

Hypothesis 7: Judge confidence in his/her final estimate will be positively related to consensus between the initial estimate and advice.

CHAPTER 2: METHODS

Participants

Judges (N = 100) were undergraduate students from a mid-sized southeastern university. Participants were mostly female (64%) and Caucasian (90%). Participants were primarily freshmen (62%) or sophomores (24%), and had an average age of 18.80 (SD = .99).

The students participated to fulfill an undergraduate research credit requirement, but were told that they could earn a monetary bonus based on the accuracy of their estimates on a knowledge quiz. The topic was U.S. History, but participants were not told this during recruitment to avoid sampling bias.

Measures

U.S. History knowledge was assessed via a five-item self report scale written by the author for this study (e.g. "I consider myself something of a U.S. History buff"; see Appendix A). The reliability of this scale was excellent (α = .88). Agreeableness was measured with a ten-item scale from the international personality item pool (e.g. "I am interested in people"; see Appendix B). This scale yielded a strong reliability coefficient for my sample (α = .82), which is consistent with the reliability of this scale in other samples (Goldberg, 1992).

The U.S. History questions used for the decision-making section of this study were selected from a pool of 30 items from an earlier advice study (Gino, 2008). These questions were selected because the historical events are obscure enough that it would be unusual for participants to know the exact date, but possible for them to make a reasonable estimate. A pilot test (N = 11) was conducted to generate advice for the full study, and to narrow the test to a 15-item subset (see Appendix C). Pilot test participants received a $.50 bonus for each question they got right (out of 30). The 15 final items were selected by the researcher for roughly homogeneous variance, as well as the more subjective evaluation that decision makers would have a difficult time setting a precise date for the event. This was desirable in order to make the quality of the advice less readily apparent, particularly during the expert advisor condition, and thus make the participant suspicious of the experimental manipulation.

Accuracy was measured as the absolute difference between the judge's estimate and the correct answer. Judge confidence was measured by the range of a 90% credibility interval around their estimated date. This credibility interval will be explained further in the procedure section. Advice value was rated after each final estimate, and was measured by a single question ("How valuable was this advice to you?").

Procedure

Participants were run individually, and the experimental procedure was conducted via pencil and paper tests. Prior to beginning the decision task, participants were asked to fill out a questionnaire which included demographic questions as well as self-report measures of U.S. History knowledge and agreeableness. After completing the questionnaire, participants were informed that they would receive $0.50 every time they were within ten years of the correct answer when estimating a U.S. Historical date. Accordingly, the participants had a financial incentive for high performance, and it was in their best interest to utilize the estimates given them to make accurate judgments.

The experiment was conducted in two stages. In the first stage, participants estimated the year of 15 U.S. historical events since 1600 (e.g. "In what year was the first US satellite in orbit?"). They were informed that they could earn up to $15 based on their performance. Specifically, they would receive a $.50 bonus every time they correctly estimated the correct answer (+/- 10 years). Participants received this bonus at the conclusion of the study, and were not given feedback on the accuracy of any particular estimate.

After they had made their initial estimate, they were asked to provide lower and upper bounds for a 90% credibility interval. It was explained to the participants that they should select the lower and upper bounds so as to be 90% sure that the true answer fell somewhere between these two values. The range of this credibility interval therefore serves as a measure of item-level initial confidence.

In the 2nd phase of the experiment, participants were told that they would have the opportunity to answer these questions a second time, but would have the advantage of receiving advice beforehand. The participants were told that the advice consisted of randomly selected answers from other undergraduate students. This is an acceptable way of generating estimates in a manner that makes it difficult for the judge to predict the quality of advice across items (Gino, 2008; Yaniv & Kleinberger, 2000).

Advisor characteristics were experimentally manipulated in terms of advisor expertise and advice cost. Participants were assigned to conditions according to a random number generator. Advisor expertise had two conditions. In the 'expert advice' condition, the participants were told that the advice was randomly selected from a pool of history majors currently enrolled in an upper-level U.S. History course at Clemson. In the ‘novice advice’ condition, the participants were told that the advice was randomly selected from a pool of students that are “on average, about as qualified as you are.” However, all advice was generated in the same way, and was held constant across all novice and expert conditions.

The participants were also randomly selected to be in one of three advice cost conditions. In the ‘free’ condition, participants did not have to pay for the advice. In the two ‘costly’ conditions, judges were told; “The cost of this advice is $3. The experimenter is going to toss a coin. If the result of the coin toss is TAILS then $3 will be subtracted from your earnings. If the result of the coin toss is HEADS you will receive the advice for free.” After the coin flip, participants were reminded, “Because the coin flip was TAILS (HEADS), you (do not) have to pay $3 for this advice.”

After the coin flip (if applicable), judges were given a piece of paper with the advisor’s estimates for each question, and were also able to view their initial estimates and 90% C.I. Advice was chosen by randomly selecting an estimate from one of the pilot test participants for each question. Although the advice itself was randomly generated, participants across all conditions received the same advisor estimate. The questions for stage 2 were the same questions that had been on the first quiz, and were presented in the same order. Participants then made a final estimate, and again picked lower and upper bounds for a 90% C.I. for each question, as well as rating the value of the advice. The participants were thanked, debriefed, and dismissed.

Payment

Judges received payment according to the accuracy of their judgments. For each response that was within 10 years of the true date, they received $.50. Students in the paid advice condition were not actually charged $3 for their advice in order to maintain payment equity across conditions. Participants on average earned $4.23 for participating (SD=$1.53). All participants earned at least $1.50.

Coding

Past judgment studies have used the "weight of advice" (WOA) measure to assess advice use. WOA = |final estimate – initial estimate|/|advice– initial estimate|. If a judge changes his final estimate to the exact value of the advisor's estimate, WOA is equivalent to one. If the participant completely discounts the advice and retains her initial estimate for her final judgment, WOA for that decision would be zero. If a participant chooses the midpoint between the advisor's estimate and his own estimate, WOA would be 0.5. It is worth noting that choosing the midpoint between estimates is the optimal strategy if the judge and advisor are equally qualified and well-informed on the decision. If the judge's initial estimates are generally more accurate than the advisor, optimal weight of advice is somewhere between 0 and 0.50. If the advisor is more accurate than the judge, the optimal WOA lies between 0.50 and 1.0. This averaging strategy tends to produce more accurate answers because it cancels out error (Larrick & Soll, 2006).

Accuracy was defined as the absolute difference between the judge's estimate and the true date, such that smaller values denote more accurate answers (Accuracy = |Estimated Year – Correct Year|). Confidence was operationalized as the range between upper and lower bounds on the 90% C.I. With both accuracy and confidence, lower values signify a more accurate/more confident response. ‘Consensus’ was operationalized as the absolute difference between initial estimates and the advisor’s estimate for each question.

CHAPTER 3: RESULTS

Data Screening

A between subjects ANOVA power analysis was conducted to determine the sample size require to attain power of .80, presuming a medium effect size (r = .25). Prior research has found a medium to strong effect size for advice cost, so this is expected to be a conservative estimate of power. Results indicated that a balanced ANOVA with a sample size of 93 would be sufficient for all main effects (no between subjects interactions were specified). Due to random assignment issues, we ended up recruiting 100 participants. Frequency data for experimental condition can be found in Table 1.

All data were screened for univariate outliers. All typographical errors were checked against the original data collection packet and corrected. Outliers that were +/- 3 standard deviations from the mean were identified. No outliers were identified for any of the scales, and less than 2% of the estimates and confidence intervals were considered outliers (See Table 3). Analyses were run both including and excluding these outliers, but all results reported in this paper are with all data included unless otherwise noted. One participant did not follow directions for the estimates and confidence intervals, and was excluded from the analyses. Three participants did not indicate their gender, and a tracking error caused us to lose advice cost information on four of the participants.

It was necessary to correct for some irrational weight of advice (WOA) estimates.[4] This occurred in about 8% of the cases for this sample. All analyses were

[4] If a judge's final estimate does not fall between the advisor's estimate and their initial estimate, it is possible for WOA to be calculated as a value greater than one. Using this value would not be a rational interpretation of the data, as it is not sensible to consider the judge to exhibit a shift towards an advisor's opinion in excess of 100%.

therefore run twice. In the first set, all estimates with WOA > 1 were reset to 1, and in the second set of analyses, estimates with WOA > 1 were deleted. Although there were some slight differences in the strength of correlations, significance and directions of relationships did not change for any of the hypotheses.

Descriptive statistics for all continuous variables are listed in Table 2 and a correlation matrix can be found in Table 3. All results presented are with outliers included. In order to be consistent with prior research, all results reported here are with high WOA estimate set to a value of one.

Tests of hypotheses

To begin, it is important to establish that receiving advice actually improved decision accuracy, regardless of decision strategy. Hypothesis 1 predicted that final estimates would be more accurate than initial estimates. This hypothesis was fully supported. The participants made final estimates with substantially lower mean absolute error than their initial estimates ($t = 7.04$, $p < .05$; mean error of 25.43 years vs. 39.22 years).

Both parts of Hypothesis 2 were supported. Self Reported US History Knowledge is a significant predictor of error in a judge's initial estimate (EIE; $r = -.40$). Furthermore, US History Knowledge is significantly related to weight of advice (WOA; $r = -.41$). Agreeableness was not significantly related to WOA ($r = -.10$), which means that Hypothesis 3 was not supported.

Next, I tested whether advisor characteristics influenced advice use. As mentioned in my hypotheses, both advisor expertise (H4a) and advice cost (H5a-b) were expected to predict advice use. I conducted a series of between-subject ANOVAs with average WOA as the DV and advisor condition as the categorical IV. Since self-ratings of prior knowledge were related to WOA (see hypothesis 2b), I included US History knowledge as a covariate in all ANOVAs. The ANOVA results revealed that advisor ‘expertise’ was a significant predictor of advice use even after controlling for judge US History knowledge, $F(3, 95) = 10.43$, $p < .05$. Expert advice was valued significantly more than novice advice WOA (Table 4), supporting Hypothesis 4a.

Next, I ran an ANOVA using advice cost as the categorical predictor. Contrary to Hypothesis 5(a-b), advice cost was not a significant predictor of advice use, ($F(4, 90) = 0.55$, ns). Pairwise analyses revealed no significant differences between any of the three advice cost conditions when controlling for judge US History knowledge (Table 5). An ANOVA was conducted with advice cost, advisor expertise, and US History knowledge all included as predictors, as well as an advice cost by advisor expertise interaction Only advisor expertise ($F(12, 82) = 10.92$, $p < .05$) and US History knowledge ($F(12, 82) = 12.63$, $p < .05$) emerged as significant predictors (Table 6). A final model was run including all possible two- and three-way interactions, but did not yield any significant results (Table 7). Cell means for all conditions are presented in Table 8.

The above procedures were repeated again with judge ratings of advice ‘value’ as the DV instead of advice use. Because prior US History knowledge was a significant predictor of ‘value’ ratings, this was controlled for in all analyses. Advisor expertise approached but did not reach significance as a predictor of value ratings ($F(3, 95) = 3.29$, $p = .07$). Overall, participants in the expert advisor condition provided slightly higher value ratings than did participants in the novice advisor condition (Table 9). Advice cost also approached significance ($F(4, 94) = 2.67$, $p = .07$). However, this relationship was in the opposite direction from the prediction in Hypothesis 5c (Table 10). Pairwise comparisons indicated a significant difference between free and paid advice, such that free advice was valued somewhat more highly than paid advice (Table 11). There was no interaction between cost and expertise (Table 12, Table 13).

Hypothesis 4c proposed that judges that received expert advice would be more confident in their final estimates. However, most variation for confidence in final estimate (CFE) was accounted for by a participant's confidence in his or her initial estimate (CIE) ($F(1, 97) = 75.59$, $p < .05$). Advisor condition failed to explain a statistically significant amount of additional variance, $F(4, 94)= 1.65$, $p=.79$, meaning that Hypothesis 4c was not supported (Table 14).

Multilevel Analyses

The design of this research study allowed us to capture multiple observations for each participant. Each judge gave 15 initial estimates, and gave 15 more estimates after being provided with advice. The data collection method has given us information on the error of the initial estimate, confidence in the initial estimate, error in the final estimate, confidence in the final estimate, weight of advice, and self-ratings of advice value for each question. This within-subjects design can result in analyses with more statistical power.

However, regression analyses require an assumption of independence. Because these measurement occasions are nested within persons, it is necessary to test this assumption of independence. The Intraclass Correlation Coefficient (ICC) is a measure of non-independence. A high ICC indicates that a higher proportion of variance is explained by differences in the average level of each person, also known as intercept variance. ICCs were calculated for all dependent variables in H6a-c and H7 (see table 15). Although the error explained by intercept variance was relatively small for most DVs, even relatively small ICCs can cause researchers to overestimate effect sizes. The researcher therefore used multilevel modeling for within-person analyses.

The first set of analyses tests the hypothesis that judge confidence in his or her initial estimate would be a significant predictor of estimate accuracy. A multilevel intercept-only model was run with initial estimate accuracy as the dependent variable. No predictors were included in order to ascertain a baseline amount of error. Next, because the difficulty of the question (as well as quality of the advice) varied across the 15 items, trial was included as a predictor. Trial contributed a 16% improvement over the null model ($F(15, 1564) = 26.43$, $p <. 05$). Next, initial estimate confidence was included as a level 1 predictor, which means that it was allowed to vary randomly both within and across individuals. A significant result would demonstrate that judges are able to accurately judge the quality of their own estimates. When initial estimate confidence and trial were included as predictors, the model jumped to a 31% improvement over the null model, with confidence in initial estimate emerging as a statistically significant predictor ($F(16, 1462) = 23.25$, $p < .05$).

A second analysis was run via the same procedure to determine if confidence in one's initial estimate was a valid predictor of final estimate accuracy as well. Trial reduced error 42% relative to the intercept-only model ($F(15, 1469) = 80.56$, $p < .05$). Confidence in initial estimate (CIE) was then included as a level 1 predictor. Judge confidence in his or her initial estimate was not a significant predictor of final estimate accuracy ($F(16, 1465) = 1.43$, ns). Therefore, Hypothesis 6a was supported for initial estimate accuracy, but not supported for final estimate accuracy.

A similar procedure was run to test hypothesis 6b. Hypothesis 6b predicted that a significant portion of variation in advice use (WOA) would be explained by judge confidence in his or her initial estimate (CIE). An intercept-only model was tested with WOA as the DV to determine the variance in the null model. Including trial as a predictor reduced this variance by about 6% ($F(15, 1418) = 6.99$, $p < .05$). Adding judge confidence as an additional level 1 predictor further reduced error, with an error reduction for the full model of about 9% ($F(16, 1416) = 54.43$, $p < .05$). These results support hypothesis 6b, suggesting that people are less interested in using an advisor's opinion if they are confident in their own response.[5]

Hypothesis 6c predicted that there would be an interaction between a level 1 predictor (judge confidence in initial estimate) and a level 2 predictor (advisor expertise). In order to test this hypothesis, I added advisor expertise to the above model, and specified an expertise by confidence interaction term. Although a main effect for advisor

[5] Note that although the slope between CIE and WOA is positive, CIE is calculated as the range of a 90% confidence interval, so larger numbers actually indicates lower confidence. These findings are therefore in the hypothesized direction, indicating that judges are less likely to use advice if they have designated a narrow range of likely values.

expertise was found (*F*(18, 1399) = 14.50, *p* < .05), the interaction was not significant (*F*(18, 1399) = .004, ns). Therefore, hypothesis 6c was not supported.

Hypothesis 7 proposes that judges are more confident when there is relatively little discrepancy between their initial estimate and the advice they receive. In the final multilevel model analysis, we tested the effect of judge-advisor consensus on judge confidence in his or her final estimate (CFE). This hypothesis was supported (*F*(16, 1461) = 72.65, *p* < .05), consensus explaining 27% of the intercept variance in CFE. A post-hoc analysis was also conducted to determine if discrepancy was a significant predictor of advice use. There is a positive relationship between discrepancy and WOA, such that judges were more likely to use advice when it was further from their initial estimate, F(1, 1431)=84.05, p<.05, r=.24.

CHAPTER 4: DISCUSSION

This section will begin by summarizing and discussing how both judge and advisor characteristics may impact the use of advice within a decision making setting. With that information providing a foundation, I will then discuss the primary contributions of the current research. A discussion of limitations and future research directions will close out this section of the paper.

Judge characteristics

The current results suggest that individuals have a reasonably accurate understanding of their knowledge of a particular topic, in this case US History (H2a). Furthermore, this assessment of their own knowledge influences the extent to which they use their initial estimates when making a final decision (H2b). This is consistent with prior research findings, which have found that judges compare their own qualifications to that of their advisor (e.g. Harvey & Fischer, 1997; Sniezek et al., 2004; Yaniv & Kleinberger, 2000). Yaniv and colleagues (e.g. Yaniv, 2004; Yaniv & Kleinberger, 2000) have suggested that egocentric bias comes from an inaccurate assessment of one's own qualifications relative to an advisor's qualifications. However, even though US History Knowledge predicted advice use, WOA averaged greater than .50 for all conditions, which suggests that egocentric bias was greatly reduced or eliminated in this setting. Possible explanations for this are discussed later in this section.

This study failed to find a relationship between agreeableness and advice utilization (H3). Bonaccio (2007) did find a link between agreeableness and advice use, so this non-significant result should be interpreted with caution. Specifically, the rationale for the agreeableness-advice use link was that highly agreeable people would be more likely to engage in consensus decision making which would allow everyone to share in final decisions. Bonaccio found that highly agreeable individuals were more likely to incorporate the advice of novice advisors into their final estimate. Agreeableness primarily influenced advice use by reducing the tendency of individuals to completely discount less useful advice. Because people tended to *overweight* advice in my study, it may be that any link between agreeableness and advice use was masked by a floor effect.

Moreover, it appears that judges re-evaluated the quality of their own estimate for each question. Judge confidence in his or her initial estimate was a significant predictor of advice use, even when their global estimate of US History Knowledge was included in the model (H6b). This suggests judges make an assessment of their own competence for specific decision tasks as well as subject domains, and that this in turn influences how much people use advice. This finding slightly extends prior research suggesting that advisor estimates of confidence influences advice use.

Advisor Characteristics

Advice use was also influenced by attributes of the advisor. Throughout this study, the actual advice remained constant across all conditions, but advisor expertise and advice cost were manipulated across subjects. Half of the participants were told that the advice they were receiving was randomly sampled from students in a US History class, while the other half were told that the advice was randomly sampled from the undergraduate student body. I hypothesized that students in the expert advice condition would be more willing to give expert advisors the benefit of the doubt when the judge's initial estimate and the advisor's estimate were discrepant (H4a). Even though the expert sample was not any better informed than either the judges or the advisors in the novice condition, judges did indeed use expert advice more than novice advice. Because this was significant independent of actual advice quality, it suggests that people use advisor characteristics when they are unable to make an accurate assessment of the advice quality. It is worth noting that both novice and expert advice were weighted more heavily than the judge's own estimate. Prior research has found that people are more likely to use advice that comes from an expert (e.g. Sniezek, Schrah, & Dalal, 2004). Moreover, as in our study, this has held true even when advisor 'expertise' is simply a label and the advice itself does not vary in accuracy across conditions (Harvey & Fischer, 1997). Moreover, judges also rated expert advice as being somewhat more valuable than novice advice (H4b); although this finding was not significant, the results were in the expected direction and approached significance (p = .07). However, judges were neither more nor less confident in their final estimate after receiving 'expert' advice.

Prior research found that judges weighed advice more heavily when they had to pay for it (e.g. Gino, 2008; Patt, et al., 2006; Sniezek, et al., 2004). Surprisingly, advice cost did not influence advice use on our study. One possible explanation for this is that participants were assigned to advice conditions, and were not allowed to choose or reject advice. However, Gino (2008) found that advice cost influenced advice use even when judges were required to receive the advice. In her study, judges were receiving advice from novices that were just as qualified as themselves, yet still chose to use the advisor's opinion substantially more than their own when they had to pay for advice, thus exhibiting that sunk costs may be playing a role in advice use. My research design was built to compare sunk costs theory to other competing hypotheses for the paid advice effect. In my study, however, there was actually a slight (nonsignificant) trend to weigh advice more heavily when cost was not mentioned at all. These findings are puzzling and require follow-up research to understand what could have caused these results.

The interaction between advice cost and advisor expertise failed to reach statistical significance. Sniezek et al. (2004) found a significant interaction between pre-paying for advice and expertise. In their study, judges and advisors were live dyads (not paper persons), and expert advisors had real training in the subject matter and therefore more accurate estimates than novice advisors, and advice payment was a function of payment timing (paying pre-advice or post-advice). Judges were more likely to use expert advice if they had already paid for it, but not more likely to use novice advice that they had already purchased.

Significant Contributions

The purpose of this study was to examine how judge and advisor characteristics influence advice use. Specifically, judges vary their policy for integrating estimates according to the perceived quality of the information. This research project advanced advice research in several ways. This study tested several hypotheses that have only recently been proposed and have not been evaluated by more than one or two researchers. For example, I measured and controlled for judge characteristics which could influence advice use, such as task specific knowledge, agreeableness and item-level confidence.

This study parsed advisor characteristics from the quality of the advice itself. To my knowledge, there has only been one published research article that experimentally manipulated advisor expertise while holding advice quality constant (Harvey & Fischer, 1997). People who claim to be experts are often giving advice, even if there is little evidence that their advice is any better than a layperson's (e.g. stockbrokers), or even that they are true experts (e.g. infomercials).

Only a handful of studies have looked at the paid-advice effect at all (Gino, 2008; Patt et al., 2006; Sniezek et al., 2004). This research adds to the body of empirical studies in this domain, and raises questions about the generalizability of some of the findings. It appears that the paid-advice effect may not be a factor across all circumstances. Gino (2008) made a strong case that sunk costs sufficiently explain the paid advice effect. In Gino's study, the sunk costs were experimentally manipulated, while other theories were tested by a self-report scale. Accordingly, my study experimentally manipulated price as a signal of advice quality as well as sunk costs. I added a second control group to test a competing hypothesis, namely that price operates as a signal of advice quality. Although the results were nonsignificant, it could have revealed an important distinction between receiving 'expensive' advice for free, and receiving free advice with no price tag attached to it.

Additional studies may shed more light on this domain. For example, there are still many questions regarding advice payment. Prior research has found that payment timing can moderate the relationship between advice cost and advice use (Patt, et al., 2006). I have already launched a follow up study which mimics the design of the current study but changes the delivery of the advice and salience of the advice cost. Participants receive a show up fee in an envelope before beginning the study, and any advice costs are taken out of this envelope. Moreover, they are not given the advice packet until it has already been determined if they will need to pay for the advice. If this study were to find results in the predicted direction, we may conclude that the order which judges pay for and receive advice is an important factor indeed.

Limitations

The experimental design of this research design allowed for a great deal of control. In particular, advisor characteristics were held constant across participants within each condition. This boosted the internal validity of the study, but an experimental design like this has relatively low external validity. Furthermore, there was only a relatively small monetary incentive at stake for the decision task. It is important to replicate these findings outside of the laboratory, and demonstrate that these tendencies generalize to important decision tasks as well as trivial ones.

Another limitation of this study is that judges were unable to form a relationship with the advisor across time. Advisor trust and reputation formation in the real world is a longitudinal process, although judges can make an assessment of advice quality relatively quickly (Yaniv & Kleinberger, 2000). The research study was relatively information-impoverished, allowing judges to only focus on the characteristics of interest to this study. As mentioned earlier, all participants answered the same questions and received the same advice, regardless of condition. It is possible that it would be more informative to vary task characteristics such as task difficulty as well as advice quality characteristics.

There are several limitations of this study which may explain why advice cost was not a significant predictor of advice use. It is possible that sunk costs didn't take hold because participants did not receive a show-up fee, and thus had not received any funds at the point that they had money deducted. In fact, multiple participants expressed concern that they didn't even earn enough money to pay for the advice. Another limitation of this study is that participants did not choose whether or not they wanted to receive the advice. Prior research (Gino, 2008; Study 2) found that the paid advice effect was robust enough to operate even without judges choosing to receive advice. However, it is possible that this does not generalize across all situations. In addition to this, participants were randomly assigned to the paid advice and costly no payment conditions via a coin flip. Some participants may have resented the fee for this unsolicited advice. The combination of these two manipulations may have detached the payment from the advice itself. One could even speculate that the research procedure may have led judges to feel like they had relatively little control over their environment. Future research is envisioned in which the parameters that influence the paid advice effect are explored more fully.

Future Research Directions

It would be interesting to explore research on other individual differences that can influence judge-advisor perceptions. Yaniv and Kleinberger (2000) found that participants are sensitive to manipulations in advice quality, but the advice itself is just one cue that can signal advice quality. More broadly, this could be tied to Brunswick's lens model, where judges form impressions of various cues that could predict advice quality. Feng and MacGeorge (2006) identified a host of advisor characteristics that may influence advisor trust and advice use. It may be fruitful to conduct further experimental studies which manipulate characteristics of the advice and the advisor.

It is important to explore individual differences in judges which may influence the advice process. Although we did not find agreeableness to be a significant predictor of advice use, there may be other traits which make people more or less willing to seek and use advice. This environment was a relatively weak situational cue for interpersonal behavior, since we were using papers persons. Furthermore, it may be feasible to experimentally manipulate judge characteristics, such as perceptions of his or her competency. Recent performance may outweigh the sobriety of a long-term outlook, and learning and memory biases may prevent persons from having an accurate view of their own ability. If judges make inaccurately optimistic or gloomy predictions of their task efficacy, this will probably influence both advice seeking behavior and advice use.

One research area that we did not explore was the influence of situational characteristics on advice use. Task difficulty has been shown to have a profound impact on WOA (Gino & Moore, 2007). People generally expect themselves to do better than average on very easy tasks, and worse than average on very difficult tasks. The task that we had participants perform was relatively difficult, which may explain why the judges weighted the advice so heavily. However, there are many task dimensions that have not yet been explored in detail, such as the nature of the task, task length and task complexity.

Relatively little advice research has explored interactions across multiple sources. For example, there are only a few studies which have manipulated judge expertise and advisor expertise simultaneously (e.g. Harvey & Fischer, 1997). Prior research has shown that advice cost can influence advice use even if advisors do not get to choose whether or not they receive the advice (Gino, 2008), yet this may not be true in all situations. Moreover, it may be profitable to examine how individuals integrate advice (e.g. Schrah, 2006). Advice literature may benefit from a more harmonious integration with multiple cue probability learning (MCPL) research, since judges often use information acquired both from other decision makers and from more passive sources.

Finally, advice cost research has focused on purchasing advice in blocks, before seeing the details of the specific advice task (e.g. Gino, 2008). However, in many situations, judges are able to form their own evaluation before deciding whether or not to pay for help. For example, a homeowner may evaluate his own problem-solving capabilities before paying for an electrician, or a citizen may attempt to complete his own taxes before calling a tax professional. This current research study showed that judge confidence predicted advice use both at the person level and at the decision item level. If an individual was confident in his or her initial estimate, he was much less likely to use advice. It may be profitable to conduct a study that allows participants to purchase advice on a per-question basis to see whether initial confidence will influence a judge's willingness to pay for advice.

Conclusion

This research study makes an important contribution to the advice literature by demonstrating a link between judge confidence and advice use at multiple levels of measurement. Furthermore, advisor expertise was found to be a significant predictor of advice use even when advisor quality did not vary across advisor conditions. While it is surprising that the advice cost manipulation was inconclusive, this nonsignificant finding raises important questions about the situational specificity of the paid advice effect. Future research should seek to uncover important task, judge, and advisor characteristics which may moderate this relationship.

APPENDICIES

Appendix A: Decision-specific (U.S. History) Knowledge

1. I consider myself something of a U.S. History buff
2. I know a lot about U.S. History
3. I’ve been exposed to a lot of information on U.S. History.
4. I’m more informed about U.S. History than the average person
5. I’ve learned a good bit of U.S. History

Appendix B: Agreeableness Scale

Agreeableness (IPIP Big Five, Alpha=.82)

1. I am interested in people
2. I sympathize with others' feelings
3. I have a soft heart
4. I take time out for others.
5. I feel others' emotions
6. I make people feel at ease.
7. I am not really interested in others
8. I insult people
9. I am not interested in other people's problems
10. I feel little concern for others

Appendix C: US History Questions for Pilot Test
(Gino, 2008)

1. When did the Congress declare war on Mexico?
2. In what year did Vietnam fall to Communists?
3. In what year was the Korean armistice signed?
4. In what year was NATO formed?
5. In what year was the Cuban missile crisis?
6. In what year did the Pilgrims reach Cape Cod?
7. When was the Truman Doctrine announced?
8. When was the Berlin wall built?
9. In what year was OPA (Office of Price Administration) established?
10. In what year was the first transcontinental railroad completed?
11. In what year the Women's rights convention at Seneca Falls take place?
12. In what year was the Presidential Succession Act?
13. When did Texas declare its independence (Battle of the Alamo)?
14. In what year was the National Labor Union formed?
15. In what year was the first US satellite in orbit?
16. When was the Standard Oil Trust organized?
17. In what year was the Civil Rights Act?
18. In what year was the Voting Rights Act?
19. When did the first American astronaut orbit earth?
20. In what year was SEC (Securities Exchange Commission) created?
21. When was the Emancipation Proclamation?
22. When was the Panama Canal opened to shipping?
23. In what year was Louisiana purchased?
24. When was the Haymarket Riot?
25. When did Mussolini seize power in Italy?
26. When was the Massachusetts Bay Colony founded?
27. When did the Korean War start?
28. In what year were US troops sent to South Vietnam?
29. When was the American Constitution first drafted?
30. When was the Bill of Rights ratified?

Appendix D: US History Questions used in full study

1. In what year was the first transcontinental railroad completed?
2. In what year was the first US satellite in orbit?
3. In what year was NATO formed?
4. When was the Massachusetts Bay Colony founded?
5. When was the Berlin wall built?
6. In what year was OPA (Office of Price Administration) established?
7. When did Texas declare its independence (Battle of the Alamo)?
8. When did Mussolini seize power in Italy?
9. When was the Truman Doctrine announced?
10. When did the first American astronaut orbit earth?
11. In what year was SEC (Securities Exchange Commission) created?
12. When was the Panama Canal opened to shipping?
13. In what year was Louisiana purchased?
14. When was the Haymarket Riot?
15. In what year was the Korean armistice signed?

Table 1

Between-Subjects distribution table

Cost/Expertise	Free	CNP*	Paid	Total
Novice	14	16	17	47
Expert	16	16	15	47
Total	30	32	32	94

*Costly No Payment condition. Participants were told they might have to pay for advice but a coin flip determined that they would not actually have to pay for the advice.

Table 2: Descriptive Statistics

Variable Name	N	Mean	Std. Deviation
Age	100	18.88	.99
Gender	97	.64	.48
Class	100	1.59	.90
Agreeableness	100	5.77	.69
US History Knowledge	100	3.20	1.31
Error in Initial Estimate (EIE)	99	39.22	20.39
Error in Final Estimate (EFE)	99	25.43	4.41
Confidence in Initial Estimate (CIE)	99	51.52	43.32
Confidence in Final Estimate (CFE)	99	40.34	17.26
Weight of Advice (WOA)	99	.67	.19
Advice Value	99	4.76	1.01

Table 3

Correlation Matrix

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Age	-										
2. Class	.88*	-									
3. Gender	.05	.08	-								
4.Agreeableness	.10	.08	.31*	.82							
5. U.S. History	.01	.07	-.21*	.00	.88						
6. Avg. Value	.00	.00	.07	.04	-.30*	.86					
7. Avg. EIE	-.09	-.08	.18	.09	-.40*	.10	.65				
8. Avg. EFE	-.15	-.13	.04	.15	-.12	.15	.31*	.03			
9. Avg. CIE	.14	.16	-.06	-.11	-.30*	.14	.54*	.10	.84		
10. Avg. CFE	.10	.14	.06	-.20*	-.28*	.26*	.25*	.06	.66*	.83	
11. Avg. WOA	.03	-.01	.01	-.10	.41*	.50*	.43*	.26*	.31*	.23*	.82

*p<.05

**EIE = Error in Initial Estimate, EFE = Error in Final Estimate, CIE = Confidence range in Initial Estimate, CFE = Confidence range in Final Estimate, WOA = Weight of Advice, Avg. Value = judge ratings of advice value for each advisor estimate

Table 4:

ANOVA Table for Advisor Expertise, controlling for US History Knowledge

	SS	df	MS	F value	Sig.	r	r-square
Corrected Model	.881	2	.440	16.31	.000	.51	.26
Intercept	43.871	1	43.871	1624.42	.000		
US History	.692	1	.692	25.64	.000	.44	.20
Expertise	.282	1	.282	10.43	.002	.28	.08
Error	2.566	95	.027				
Total	47.425	98					
Corrected Total	3.446	97					

DV= Weight of Advice

Table 5:

ANOVA Table for Advice Cost, controlling for US History Knowledge

	SS	df	MS	F value	Sig.	R	r-square
Corrected Model	.504	3	.168	5.61	.001	.40	.16
Intercept	42.144	1	42.144	1407.82	.000		
US History	.451	1	.451	15.07	.000	.38	.14
Advice Cost	.033	2	.016	.55	.581	.10	.01
Error	2.694	90	.030				
Total	45.624	94					
Corrected Total	3.198	93					

DV= Weight of Advice

Table 6:

Full WOA model (all 2- and 3-way interactions included)

	SS	df	MS	F value	Sig.	R	r-square
Corrected Model	.810	6	.135	4.92	.000	.50	.25
Intercept	42.267	1	42.267	1539.92	.000		
Advice Cost	.026	2	.013	.48	.620	.09	.01
Expertise	.286	1	.286	10.44	.002	.30	.09
US History	.475	1	.475	17.32	.000	0.39	.15
Cost*Expertise	.017	2	.008	.30	.740	.07	.01
Error	2.388	87	.027				
Total	45.624	94					
Corrected Total	3.198	93					

Table 7: WOA values for Advice Cost and Advisor Expertise

	SS	df	MS	F value	Sig.	r	r-square
Corrected Model	.883	6	.08	2.84	.00	.53	.28
Intercept	40.492	1	40.492	1434.35	.00		
Expertise	.308	2	.308	10.93	.00	.31	.10
Advice Cost	.019	1	.010	.34	.71	.08	.01
US History	.357	1	.357	12.63	.00	.33	.11
Cost*Expertise	.013	2	.006	.22	.80	.07	.01
Cost*US History	.031	2	.015	.54	.58	.10	.01
Expertise*History	.019	1	.019	.68	.41	.08	.01
Expertise*Cost*History	.029	2	.014	.51	.60	.10	.01
Error	2.315	82	.028				
Total	45.624	94					
Corrected Total	3.198	93					

Table 8

WOA cell means by condition

	Free Advice	CNP Advice	Paid Advice	Total
Novice Advisor	0.65	0.62	0.58	0.62
Expert Advisor	0.74	0.72	0.73	0.72
Total	0.70	0.67	0.65	

Table 9:

ANOVA Table for Advisor Expertise, controlling for US History Knowledge

	SS	df	MS	F value	Sig.	r	r-square
Corrected Model	12.172	2	6.086	6.68	.002	.35	.12
Intercept	2227.42	1	2227.42	2447.47	.000		
Expertise	2.99	1	2.99	3.29	.073	.17	.03
US History	10.34	1	10.34	11.36	.001	.32	.10
Error	86.459	95	.91				
Total	2329.08	98					
Corrected Total	98.631	97					

DV=Advice Value

Table 10:

ANOVA Table for Advice Cost, controlling for US History Knowledge

	SS	df	MS	F value	Sig.	R	r-square
Corrected Model	14.790	3	4.930	5.366	.002	.39	.15
Intercept	2141.462	1	2141.462	2330.645	.000		
US History	9.427	1	9.427	10.260	.002	.31	.10
Advice Cost	4.961	2	2.480	2.699	.073	.23	.05
Error	82.695	90	.919				
Total	2247.240	94					
Corrected Total	97.485	93					

DV= Advice Value

Table 11:

ANOVA Table for Advice Cost and Value, controlling for US History Knowledge

Hypothesis 5a-b	Value	Std. Error
Free Advice	5.02	.176
CNP Advice	4.86	.170
Paid Advice	4.48	.169

Table 12:

Advice Cost and Advisor Expertise, controlling for US History knowledge

	SS	df	MS	F value	Sig.	R	r-square
Corrected Model	17.774(a)	6	2.962	3.233	.006	.43	.18
Intercept	2139.720	1	2139.720	2335.386	.000		
Expertise	2.249	1	2.249	2.455	.121	.15	.02
Advice Cost	4.819	2	2.409	2.630	.078	.22	.05
US History	10.206	1	10.206	11.139	.001	.32	.10
Cost*Expertise	.700	2	.350	.382	.684	.08	.01
Error	79.711	87	.916				
Total	2247.240	94					
Corrected Total	97.485	93					

DV=Advice Value

Table 13:

Value ratings by condition, controlling for US History knowledge

	Free Advice	CNP Advice	Paid Advice
Novice Advisor	4.93	4.58	4.39
Expert Advisor	5.12	5.14	4.58

Table 14:
CFE by expertise, controlling for CIE

	SS	df	MS	F value	Sig.	R	r-square
Corrected Model	12882.241(a)	2	6441.121	37.471	.000	.66	.44
Intercept	28122.639	1	28122.639	163.602	.000		
Expertise	12.466	1	12.466	.073	.788	.00	.00
Confidence in Initial Estimate	12582.799	1	12582.799	73.200	.000	.66	.43
Error	16330.208	95	171.897				
Total	188854.161	98					
Corrected Total	29212.449	97					

DV= Confidence in Final Estimate

Table 15

Intraclass Correlation Coefficients

Variable	ICC
Weight of Advice (WOA)	.19
Error in Initial Estimate (EIE)	.08
Confidence in Initial Estimate (CIE)	.39
Error in Final Estimate (EFE)	.08
Confidence in Final Estimate (CFE)	.11

REFERENCES

Arkes, H. R., & Ayton, P. (1999). The sunk cost and concorde effects: Are humans less rational than lower animals? *Psychological Bulletin, 125*(5), 591-600.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124-140.

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior & Human Decision Processes*, *39*(1), 133.

Arkes, H. R., & Hutzel, L. (2000). The role of probability of success estimates in the sunk cost effect. *Journal of Behavioral Decision Making, 13*(3), 295-306.

Bogle, J. C. (1999). *Common sense on mutual funds: New imperatives for the intelligent investor*. New York: Wiley.

Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: The Free Press.

Feng, B., & MacGeorge, E. L. (2006). Predicting receptiveness to advice: Characteristics of the problem, the advice-giver, and the recipient. *Southern Communication Journal*, *71*(1), 67-85.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology*, *75*(6), 728-731.

Gino, F. (2008). Do we listen to advice just because we paid for it? the impact of advice cost on its use. *Organizational Behavior and Human Decision Processes*, *107*(2), 234-245.

Gino, F., & Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making, 20*(1), 21-35.

Godek, J., & Murray, K. B. (2008). Willingness to pay for advice: The role of rational and experiential processing. *Organizational Behavior & Human Decision Processes*, *106*(1), 77-87.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26-42.

Gross, J.J., & Oliver, P.J. (1995). Facets of emotional expressivity: three self-report factors and their correlates. *Person individ diff,* 9, 555-568.

Harvey, N., Harries, C., & Fischer, I. (2000). Using advice and assessing its quality. *Organizational Behavior & Human Decision Processes, 81*(2), 252-273.

Kolb, J. A. (1999). The effect of gender role, attitude toward leadership, and self-confidence on leader emergence: Implications for leadership development. *Human Resource Development Quarterly, 10*(4), 305-320.

Kruger, J. (1999). LakeWobegon be gone! The ‘‘below-average effect’’ and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77*(2), 221–232.

Larrick, R. P., & Soll, J. B. (2006). Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science, 52*(1), 111-127.

Lichtenstein, S., & Fischoff, B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance,* 1977, 20, 159-183.

Lichtenstein, S., Fischoff, B., & Phillips, L.D. (1982). Calibration of probabilities, The state of the art to 1980. In D. Khaneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

Micklethwait, J., & Wooldridge, A. (1996). Have you consulted a witch doctor lately? *Forbes, 158*(12), 204-211.

Monroe, K. B. (1973). Buyers' subjective perceptions of price. *Journal of Marketing Research, 10*(1), 70-80.

Moore, D. A., & Kim, T. G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology, 85*(6), 1121–1135.

Ordóñez, L. D. (1998). The effect of correlation between price and quality on consumer choice. *Organizational Behavior & Human Decision Processes, 75*(3), 258-273.

Patt, A. G., Bowles, H. R., & Cash, D. W. (2006). Mechanisms for enhancing the credibility of an adviser: Prepayment and aligned incentives. *Journal of Behavioral Decision Making, 19*(4), 347-359.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534-552.

Pfeffer, J., & Sutton, R. I. (2006). *Hard facts, dangerous half-truths, and total nonsense*. Boston, MA: Harvard Business School Press.

Schrah, G. E., Dalal, R. S., & Sniezek, J. A. (2006). No decision-maker is an island: Integrating expert advice with information acquisition. *Journal of Behavioral Decision Making, 19*(1), 43-60.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis*, *13*(6), 675-682.

Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidence in group decision making. *Organizational Behavior and Human Decision Processes, 52,* 124–155.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge--advisor decision making. *Organizational Behavior & Human Decision Processes*, *62*(2), 159-174.

Sniezek, J. A., Schrah, G. E., & Dalal, R. S. (2004). Improving judgement with prepaid expert advice. *Journal of Behavioral Decision Making, 17*(3), 173-190.

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge-advisor system. *Organizational Behavior and Human Decision Processes, 84*(2), 288-307.

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 & Cognition, 35*(3), 780-805.

Trafimow, D., & Sniezek, J. A. (1994). Perceived expertise and its effect on confidence. *Organizational Behavior & Human Decision Processes, 57*(2), 290-302.

Tversky, A., & Kahneman, D. (1986). Judgment under uncertainty: Heuristics and biases. In H. R. Arkes & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader.* (pp. 38-55). New York, NY US: Cambridge University Press.

Van Swol, L. M., & Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology, 44*(3), 443-461.

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

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*(2), 558-563.

Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General, 124*(4), 424-432.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes, 83*(2), 260-281.

Yates, J. F., Price, P. C., Ju-Whei Lee, P. C., & Ramirez, J. (1996). Good probabilistic forecasters: The 'consumer's' perspective. *International Journal of Forecasting, 12*(1), 41-56.

Zarnoth, P., & Sniezek, J. A. (1997). The social influence of confidence in group decision making. *Journal of Experimental Social Psychology, 33*(4), 345-366.