University of Calgary

PRISM Repository **https://prism.ucalgary.ca**

The Vault Open Theses and Dissertations

2013-08-20

Experimental Studies on Motivation and Performar

Hendijani, Rosa

Hendijani, R. (2013). Experimental Studies on Motivation and Performance (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca. doi:10.11575/PRISM/26066 http://hdl.handle.net/11023/880

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Experimental Studies on Motivation and Performance

by

Rosa Hendijani

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

HASKAYNE SCHOOL OF BUSINESS

CALGARY, ALBERTA

AUGUST, 2013

© Rosa Hendijani 2013

Abstract

This dissertation consists of three experimental studies to test the effect of motivational factors, such as monetary rewards and social relationships, on performance in different contexts, such as health care. The three studies can be thought of as a single body of work because they use the same methodology to address different research questions related to motivation and performance. In each study, we designed an experiment to test our hypotheses that were developed based on the related literature on motivation in various contexts, including health care.

The first study examined the effect of external reward and intrinsic motivation on overall task motivation and performance, using a laboratory experiment; we used self-selection into an area of knowledge for assessing intrinsic motivation. The results of this experiment provided support for our hypotheses regarding the positive effect of external reward and intrinsic motivation on overall task motivation and performance.

In the second and third studies, we focused on motivational factors affecting referral processes in health care systems. The second study examined the effect of social relationships on referral rate. Using an online survey, we examined how a shift from a decentralized referral system, characterized by close relationship between general practitioners and specialists, to a centralized system, in which there is no relationship between the referring general practitioner and the specialist, would affect referral decision making. We found partial support for our hypothesis on the effect of social relationships and referral rate. Medical doctors who had high confidence in their referral decision making referred significantly fewer patients under the close relationship condition in comparison with the centralized referral system.

The third study examined the effect of fundholding and pay-for-performance schemes on referral rate and referral appropriateness. While we could not find significant statistical support

for our hypotheses, the results were in the direction that we predicted. Both fundholding and pay-for-performance schemes decreased referral rate in comparison with the fixed pay treatment. In addition, pay-for-performance resulted in more appropriate referrals in comparison with fundholding and fixed pay schemes.

The dissertation chapters are in the following order: chapter one gives an introduction to the field of motivation and performance. Chapter two provides a literature review of intrinsic-extrinsic motivation. Chapter three presents our first study on the effect of external reward and intrinsic motivation on overall task motivation and performance. Chapter four provides a literature review on the factors affecting referral decision making in healthcare systems. Chapters five and six present the second and third studies on the effect of social relationships and financial schemes on referral patterns respectively. Finally, chapter seven provides concluding remarks regarding the results of our three studies and future research directions.

Acknowledgements

It is with immense gratitude that I acknowledge the support and help of my supervisor, Dr. Diane Bischak, who assisted me in all the steps of my doctoral research from study design to implementation and analysis. Without her thoughtful guidance and her critical remarks, this research would not become possible. She taught me the greatest essence of a scientific discourse: to think critically and to interpret impartially.

I would also like to thank Dr. Mehdi Mourali and Dr. Piers Steel, my committee members, for their valuable comments on the design and implementation of my experiments and their guidance during my doctoral research. In addition, I express gratitude to Dr. Piers Steel for his constructive recommendations on the direction of my future research.

Finally, I would like to thank Dr. Joseph Arvai, my previous co-supervisor, whose creative ideas and kind support assisted me in shaping the direction of my doctoral research.

I would like to dedicate this dissertation to my parents who cordially supported me during all years of my education.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction

One of the most important aspects of managing any organization is motivating individuals to perform in the most effective and efficient ways. Motivation is one of the critical determinants of individuals' performance in a variety of organizational settings. Organizations use different types of motivational mechanisms, such as feedback or external rewards, to improve employee performance.

Since the eighteenth century, motivation has been widely studied in the psychological sciences. According to the social psychology literature, the term "motivation" describes why a person chooses one response over another or makes a specific response (Fiske et al., 2010). Similar to social psychologists, behavioral operations management researchers have addressed the issue of motivation through studying "intentions." Intentions refer to the actual goals or underlying motivational mechanisms of decision makers that result in their actions. While effective and efficient actions of decision makers are ultimately the focus in operations management, studying the underlying basis of the motivations of those decision makers can provide insights into why operational policies sometimes have unintended or unexpected effects in practice.

Individuals' motivation can be due to their personal preferences or due to a variety of external factors (Bendoly et al., 2006). Different types of motivation can be distinguished on the basis of the reasons or goals that give rise to an action (Ryan & Deci, 2000). Based on self-determination theory, the most basic distinction is between intrinsic motivation and extrinsic motivation (external reward). Intrinsic motivation can be defined as “the doing of an activity for its inherent satisfaction rather than for some separable consequences” (Ryan & Deci, 2000).

Extrinsic motivation, on the other hand, is external to the activity and refers to doing something for the purpose of achieving a separable desired outcome (Ryan & Deci, 2000).

Motivational drivers can be positioned along the continuum between extrinsic and intrinsic motivations depending on the perceived source of regulation (i.e., controlled versus self-determined) (Deci & Ryan, 2000). These motivational factors range from the case of "pure" extrinsic motivation such as, monetary rewards or feedback, at one extreme to more autonomous forms of extrinsic motivation, such as status seeking and social relationship, to pure forms of intrinsic motivation, which includes doing an activity for its inherent interest. Previous studies have consistently found that positive feedback enhances motivation and performance (Deci et al., 1999; Nadler, 1979). In contrast to the effects of positive feedback, the exact effect of different types of monetary rewards on motivation and performance has not been consistent in the literature (Deci et al., 1999) and needs further examination.

In this dissertation, we examine the effect of diverse types of motivational factors on performance in different contexts. Among different types of motivations, we specifically study the effect of pure intrinsic motivation, social relationships, and monetary rewards as three motivational drivers of performance. We designed three experimental studies to test the effect of these motivational factors on performance in different contexts, such as health care. In each study, we designed an experiment to test our hypotheses that were developed based on the related literature on motivation in various contexts, including health care. The three studies can be thought of as a single body of work because they use the same methodology to address different research questions related to motivation and performance.

The next chapters of the dissertation are in the following order: in chapter two, we provide a literature review of intrinsic-extrinsic motivation. In chapter three, we present our first study on the effect of external reward and intrinsic motivation on overall task motivation and performance. Chapter four provides a literature review on the factors affecting referral decision making in healthcare systems. Chapters five and six present the second and third studies on the effect of social relationships and financial schemes on referral patterns respectively. Finally, chapter seven provides concluding remarks regarding the results of our three studies and future research directions.

Chapter Two: Motivation Literature

2.1 Introduction

The effect of external rewards on intrinsic motivation and performance has been widely addressed by both psychologists and economists over the past few decades. However, the studies have not reached a conclusion on the effect of different types of monetary rewards, as an important type of external rewards, on intrinsic motivation and performance. In the next section, we will first define different types of motivation using the self-determination theory approach. Then, we will examine the operationalization of the motivation constructs in the literature and see how intrinsic motivation and external rewards have been categorized and measured by the motivation scientists. Finally, we will provide related theories of motivation and their predictions on the effect of different types of external rewards on motivation and performance.

2.2 External Reward versus Intrinsic Motivation

Based on self-determination theory, different types of motivation can be distinguished based on the reasons or goals that give rise to an action (Ryan & Deci, 2000). The most basic distinction is between intrinsic motivation and extrinsic motivation (external reward). Intrinsic motivation refers to doing something that is inherently rewarding. Such inherent reward is called intrinsic motivation. Intrinsic motivation can be defined as "the doing of an activity for its inherent satisfactions rather than for some separable consequences" (Ryan & Deci, 2000). Extrinsic motivation, on the other hand, is external to the activity and refers to doing something for the purpose of achieving a separable desired outcome (Ryan & Deci, 2000).

Different motivational factors can be positioned along the continuum between extrinsic and intrinsic motivations depending on the perceived source of regulation (i.e., controlled versus

self-determined) (Deci & Ryan, 2000). These motivational factors range from the case of "pure" extrinsic motivation, such as monetary rewards or feedback, at one extreme to more autonomous forms of extrinsic motivation, such as status and opportunity, to pure forms of intrinsic motivation. In this dissertation, we focus on extrinsic and intrisic motivation in their pure form.

Until 1971, behaviorists focused only on the effect of external rewards on behavior (Skinner, 1953; 1945). When external rewards were administered subsequent to a behavior, they increased the likelihood that the behavior be repeated. This effect persisted as long as the reward was present. When rewards were eliminated, the likelihood that the behavior be repeated decreased to the prereward condition (Deci et al., 1999; Loveland & Olley, 1979).

Unlike behavioral theorists, cognitive evaluation theorists argued that some activities provide their own inherent reward. Activities that have such intrinsic interest in themselves were called “intrinsically motivated activities” (Deci, 1971; White, 1959).

2.3 Intrinsic Motivation Measurement

Studies on intrinsic motivation use two main types of measures: independent measures and dependent measures of intrinsic motivation.

2.3.1 Independent Measures of Intrinsic Motivation

Although most studies of intrinsic motivation use dependent measures of intrinsic motivation in their experiments, some studies use independent measures of intrinsic motivation in order to test the interaction effect of external reward and intrinsic motivation on dependent variables such as satisfaction and performance. These studies mainly use generally interesting versus dull activities for the manipulation of intrinsic motivation. As one example of these activities, the use of boring versus interesting puzzles can be mentioned (Calder & Staw, 1975b;

Daniel & Esser, 1980). In order to manipulate intrinsic motivation, Calder and Staw (1975b) used blank puzzles in one group (low intrinsic motivation group) and puzzles with interesting pictures in another group (high intrinsic motivation group). A manipulation check for intrinsic motivation in this study was done after the experiment was complete. Other studies also used boring versus interesting puzzles for the manipulation of intrinsic motivation (Daniel & Esser, 1980).

However, a few studies within the industrial-organizational psychology literature attempted to use tasks that are more similar to the ones in the real world settings. As an example, in Hamner and Foster's (1975) study, participants in the boring treatment were asked to code and transfer the scores from a recent math survey while participants in the interesting treatment were asked to code and transfer the scores from a recent sexual attitude survey. In another study, the task was proof reading and intrinsic motivation was manipulated by having participants read material from a law review in the boring condition and from short stories in the interesting condition (Philips & Freedman, 1985).

2.3.2 Dependent Measures of Intrinsic Motivation

Most studies in the intrinsic motivation literature focus on the effect of external reward on intrinsic motivation. In these studies, intrinsic motivation is the dependent variable of the study. Therefore, the assessment of intrinsic motivation is done subsequent to the experimental period during the reward phase (e.g., Morgan, 1981) or after the reward is terminated (e.g., Ryan et al., 1983). These studies use two main measures for operationalizing intrinsic motivation: self-reports of interest and free-choice behavior.

The primary measure of intrinsic motivation, introduced by Deci (1971), is called free-choice behavior. It refers to the degree to which participants return to and persist at the target task during a free-choice period subsequent to the experimental phase (Deci, 1971). In most studies using free-choice persistence, the assessment of the free-choice behavior was based on the amount of time spent on the target task during the free-choice period (e.g., Amabile et al., 1986, Exps. 1 and 3; Boggiano et al., 1985, 1982; Brennan & Glover, 1980). In other studies, free-choice persistence was measured by the number of trials or successes with the activity during the free-choice period (e.g., Harackiewicz et al., 1984, Exps. 2 and 3; Weiner, 1980; Pittman et al., 1977). In a few other studies, free-choice behavior was measured by the proportion of participants who spent any time with the target activity (Pallak et al., 1982; Swann & Pittman, 1977, Exps. 1 and 2).

The second measure of intrinsic motivation is self-reported interest in the target activity, which is assessed after the task either with a single item or with multiple items (e.g., Harackiewicz et al., 1984, Exps. 1, 2, and 3; Harackiewicz & Manderlink, 1984; Weinberg & Jackson, 1979). Generally, the items include questions assessing interest and enjoyment in the experimental activity.

Other measures of intrinsic motivation used in the literature include interest and satisfaction (Calder & Staw, 1975b), performance during the free time period (i.e., number of solved puzzles or problems, number of completed drawings) (Hamner & Foster, 1975), and willingness to volunteer for future activities without reward (Daniel & Esser, 1980; Calder & Staw, 1975b).

2.4 External Rewards Contingencies

The literature review on external reward shows that there are different types of external reward. Reward contingencies play an important role on the effect of external reward on intrinsic motivation and performance (Deci et al., 1999). Rewards can be categorized based on whether they are expected while a person is doing the task and, if yes, on what specific behaviors they are dependent (Ryan et al., 1983).

The first category of rewards referred to as *task-noncontingent* rewards are given for something other than engagement in the related activity. One example is the expected reward given to people for participating in an experimental session, independent of what they do in that session (Ryan et al., 1983). In this case, people are rewarded only for their presence in the session without considering the completion or quality of the task. This type of reward is analogous to hourly payments in the real world. People are paid only for being present on the job and not for any particular behavior (Ryan et al., 1983). Ryan et al. use the term task-noncontingent for these types of rewards in order to distinguish these rewards from noncontingent rewards in the context of learned helplessness theory (Seligman, 1975). Task-noncontinegnt rewards, as defined by Ryan et al. (1983), are non-contingent to task behavior. But, they are contingent on attendance, so they are controllable and predictable by the person who receives the reward and will not induce helplessness and consequent decrements in intrinsic motivation.

The second category includes *task-contingent* rewards which are given for doing or completing the related activity. For example, a person is paid a certain amount for solving each puzzle or assembling each model. Task contingent rewards are usually given for engaging in or

completing an activity without respect to quality of performance (Ryan et al., 1983). Therefore, task contingent rewards can be divided into two categories of *completion-contingent* and *engagement-contingent* rewards (Deci et al., 1999). Completion-contingent rewards are dependent on completing a target task. Engagement-contingent rewards, on the other hand, are dependent upon engaging in the activity but do not require completing it. Task-contingent rewards can be roughly compared to the piece-rate payment systems in the real world.

The third category of rewards includes *performance-contingent* rewards that are interpreted as rewards given for performing up to or above a specified level of performance, a standard, or some specific criteria (e.g., doing better than 80% of the other participants) (Ryan et al., 1983). The focus here is on the quality of the activity and whether the person is meeting some type of standard or not. Performance-contingent rewards could be considered analogous to certain types of bonus or incentive payment structures used in the real world (Ryan et al., 1983).

2.5 Theories of Motivation

2.5.1 Theories Explaining Additive Effect of Motivations

Most theories of worker motivation, especially behavioral theories, assume that intrinsic motivation and external reward are additive (Hamner & Foster, 1975; Porter & Lawler, 1968). People will be more motivated to perform an activity that combines both types of rewards than an activity where only one kind of reward is present. As examples of these types of theories, expectancy theory (Vroom, 1964), agency theory (Eisenhardt, 1989), and reinforcement theory (Skinner, 1953) can be mentioned.

2.5.1.1 Expectancy Theory

Expectancy theory was developed most prominently by the early works of Tolman (1932) and Lewin (1938). It posits that people try to maximize their expected satisfaction with outcomes (Vroom, 1964). The basic notion of the theory is that individuals have cognitive expectancies, anticipations about the consequences of a certain act, and valuations of these expectancies that may vary from strongly positive to strongly negative (Jorgenson et al., 1973). Therefore, an individual's motivation in a particular situation is a function of two factors: 1. the expectancy about the relationship between effort and a particular outcome (e.g., the expectation about the relationship between different levels of performance and payment). This relationship is called the effort-outcome expectancy. 2. The valence (attractiveness) of the outcome. These two factors result in a certain level of motivation which will lead people to choose a certain level of effort that they believe will lead to the desired outcome (Bonner & Sprinkle, 2002).

In expectancy theory conceptualization, the effect of financial rewards on effort and performance is dependent on two factors. First, financial reward plays the role of the outcome in a reward-based performance system. Financial rewards can have valence for a variety of reasons. Based on Vroom's conception, money has high valence because it is instrumental in obtaining desirable material goods. In addition, money has symbolic value due to its perceived relationship to prestige, status, and other similar factors (Furnham & Argyle, 1998; Zelizer, 1994). As a consequence, financial rewards, in general, have higher valence than no payment conditions if the expectation of receiving payment is greater than zero. Furthermore, contingent rewards may have higher valence than noncontingent rewards depending on the relative payment schedules (Pritchard et al., 1976).

Second, contingent reward systems lead to higher expectancies in comparison with noncontingent or no reward systems. This is due to the existence of stronger links among reward, effort, and performance in contingent reward systems (Locke & Latham, 1990; Pritchard et al., 1976; Jorgenson et al., 1973). Therefore, according to expectancy theory, an individual's motivation, effort, and performance are predicted to be significantly higher under a contingent reward system due to both an increased expectancy about the effort-outcome relationship and an increased (or at least no change in the) valence of the outcome. Considering the influential role of monetary rewards, expectancy theory models emphasize the establishment of performance-reward contingencies (instrumentalities) in order to increase performance. In addition, they clearly posit that performance will reach its maximum level, the closer the performance-reward contingency gets to a perfect relationship (Pritchard et al., 1976).

2.5.1.2 Reinforcement Theory

Reinforcement theory and operant conditioning indicate that external rewards can control behavior. They serve as a positive reinforcer of a behavioral act. When the rewards are administered subsequent to a behavior, they can increase the likelihood of repeating the behavior. Instrumental or operant conditioning requires the person to execute a behavioral response in order to get the reward. If the behavioral response is not executed, the person will not receive any reward. Thus, reward increases the frequency of repeating a behavior by reinforcing stimulus-response links.

This effect has proved to persist as long as the reward contingency is operative. However, when rewards are terminated, the likelihood that the behavior be repeated returns to the baseline condition. According to this theory, "a goal is nothing but a powerful external

incentive, defined as objects and events that affect an organism's behavior radically and reliably (e.g., food, sexual stimulation, money)" (Fiske et al., 2010; Bindra, 1959; Skinner, 1953). Therefore, the findings of reinforcement theory and operational conditioning highly support the use of external reward as a motivational strategy that can improve motivation and performance.

2.5.1.3 Agency Theory

Agency theory adds further explanation regarding the effects of financial rewards on effort and performance. It assumes that individuals try to maximize their utility, are boundedly rational, and have well-defined preferences that conform to the axioms of expected utility theory (Baiman, 1990, 1982; Eisenhardt, 1989). Furthermore, based on agency theory, individuals are presumed to be solely motivated by self-interest, where self-interest is described by a utility function that constraints two arguments of wealth and leisure. Individuals are presumed to have preferences to increase wealth and decrease effort (increase leisure).

Most models of economic behavior are based on agency theory which posits that individuals will exert effort on a task only if it contributes to their own economic well-being. Considering this assumption, performance-noncontingent rewards will not lead to increased effort and performance. Thus, similar to expectancy theory and reinforcement theory, agency theory suggests that external rewards can significantly improve motivation and performance because individuals want to maximize their utility by increasing their wealth.

In conclusion, like expectancy theory and reinforcement theory, agency theory indicates that financial rewards motivate individuals to exert more effort which, in turn, improves their performance. The increase in effort is due to the individual's valence or desire to receive the consequent payment.

2.5.1.4 Goal Setting Theory

According to goal setting theory, personal goals are the most important determinant of effort (Locke & Latham, 1990). People either set themselves personal goals or transform assigned goals into personal ones. Two factors of goal strength (Ajzen, 1991; Ajzen & Fishbein, 1980) and goal specificity have important effect in successful goal striving (Locke & Latham, 2006; 2002).

Empirical research shows that challenging and specific goals are more likely to be strived for and attained than moderately specific or challenging but vague goals (Locke & Latham, 2006; 2002). There are several prerequisites for this effect including frequent performance feedback, strong goal commitment, low goal complexity, and availability of required skills and means to achieve the desired goals. It is interesting to note that goal type—whether the goal is a self-set goal (chosen freely by the individual), an assigned goal (determined by others), or a goal set in interaction with others (participative goal)—does not affect success in goal striving and attainment.

Based on goal setting theory, monetary rewards can improve performance through several possible processes (Locke et al., 1981): monetary rewards may result in more spontaneous goal setting; they can lead people to set higher level, more challenging and more specific goals for themselves; and they can promote goal commitment by stimulating individuals to exert more effort to attain their desired outcome.

2.5.2 Theories Explaining the Undermining Effect of Motivations

Subsequent to the introduction of intrinsic motivation concept by Deci (1971), researchers raised the question of how external rewards would affect people's motivation for intrinsically motivated activities. Preliminary studies on the effect of external rewards on intrinsic motivation revealed that external rewards in the form of monetary incentives could undermine intrinsic motivation for interesting activities (White, 1959; Deci, 1971, 1972a, 1972b). Later studies replicated this general finding by showing that other material and symbolic rewards could also undermine intrinsic motivation (Lepper et al., 1973; Kruglanski et al., 1971). However, the undermining effect of external reward on intrinsic motivation was extended to the conclusion that external reward would undermine effort and performance (Ariely et al., 2009; Gneezy & Rustichini, 2000; Pinder, 1976; Hamner & Foster, 1975); some researchers used performance (Hamner & Foster, 1975) as a measure for intrinsic motivation in their studies. Researchers developed new theories to explain the undermining effect of external reward on intrinsic motivation. Among these theories cognitive evaluation theory, overjustification theory, and the theory of learned helplessness can be mentioned.

2.5.2.1 Cognitive Evaluation Theory

Cognitive evaluation theory is one of the most prominent theories in the undermining literature. It was primarily developed by Deci and Ryan to explain the undermining effect of external reward on intrinsic motivation (Deci & Ryan, 1985, 1980). Based on deCharms' concept of "locus of causality" (deCharms, 1968), CET asserts that intrinsic motivation consists of two main components: the need for autonomy (self-determination) and the need for competence, so the effect of external events such as reward or communication depends on how they affect

perceived autonomy (self-determination) and perceived competence. Events that satisfy the need for self-determination or competence can increase intrinsic motivation. While, those that thwart the satisfaction of the need for self-determination or competence can decrease intrinsic motivation (Deci & Ryan, 2000).

Based on this argument, rewards can have two functional aspects: informational aspect and controlling aspect. If the recipients interpret the reward as the controller of their behavior, it may thwart satisfaction of the need for autonomy, lead to a more external perceived locus of causality (deCharms, 1968) and undermine intrinsic motivation. However, if the recipients interpret the reward as an indicator of their competence, it can provide satisfaction of the need for competence and enhance intrinsic motivation.

In many cases, rewards have both informational and controlling effects that contradict each other. In such cases, these two mechanisms work against each other. Thus, additional factors such as reward contingencies must be taken into account.

According to CET, reward contingencies may affect the extent to which rewards can be interpreted as controllers of behavior versus affirmation of competence. As a result, reward contingencies can influence the degree to which different types of rewards can undermine or enhance intrinsic motivation.

Task-noncontingent rewards do not require any specific criterion for the achievement of the reward. As a result, they will have minimal controlling or informational effect on the recipient's behavior. Accordingly, these types of rewards might not have a significant undermining effect on intrinsic motivation. Engagement-contingent rewards, on the other hand, require people to work on the task to get the reward, so the reward might be perceived as

controlling the individual's behavior as regards with the task. However, since the reward does not provide competence affirmation, it is unlikely to increase perceived competence which can decrease the negative effects of perceived control. Thus, engagement-contingent rewards may undermine intrinsic motivation due to the controlling effect they have on the recipient's behavior.

Completion-contingent rewards are even more controlling because they are given only if the person completes the task. However, the receipt of completion-contingent reward may imply some level of competence if the task required skill and the individual had a sense of the specifications of good performance on the task, so the undermining effect of completion-contingent reward on performance depends on the overall interaction of controlling and informational effects. Nevertheless, CET predicted the controlling effect of completion-contingent rewards would be stronger than its competence affirming aspect, so completion-contingent rewards were predicted to undermine intrinsic motivation.

Finally, performance-contingent rewards are linked to performance. Thus, CET predicts that this type of reward will have the strongest controlling effect on people and the strongest tendency to undermine intrinsic motivation.

In summary, based on CET's predictions, task-noncontingent rewards may not undermine intrinsic motivation because they have minimal controlling or informational effect on the recipient's behavior. Engagement-contingent rewards will undermine intrinsic motivation due to the existence of controlling effect and lack of informational effect regarding individual's behavior. Completion-contingent and performance-contingent rewards may or may not undermine intrinsic motivation depending on the overall interaction of controlling and

informational effects. However, performance-contingent rewards are predicted to have the strongest tendency to undermine intrinsic motivation because they have the strongest controlling effect on performance.

Deci et al. (1999)'s meta-analysis on the effect of different types of rewards on intrinsic motivation mostly supported CET's predictions. The result of the meta-analysis showed that as predicted by CET, task-noncontingent rewards did not affect both self-reports of interest and free-choice measures of intrinsic motivation. Engagement-contingent rewards, on the contrary, undermined both self-reported intrinsic interest and free-choice intrinsic motivation. Similarly, completion-contingent rewards were found to undermine intrinsic motivation assessed with both measures of intrinsic motivation. Performance-contingent rewards had a mixed effect on intrinsic motivation. They had a negative effect on free-choice intrinsic motivation, but no significant effect on self-reports of interest (Deci et al., 1999).

2.5.2.2 Overjustification Theory

Another group of investigators provided attributional theories to interpret the undermining effect of external reward on intrinsic motivation (Higgins & Trope, 1990; Kruglanski, 1975; Lepper, 1981). The best known among these theories is the overjustification effect that is built on Bem's (1972) self-perception theory. According to this theory, people make postbehavioral attributions about the causes of their own behavior based on the behavior and the circumstances within which it occurred. When people receive a reward for doing an activity, it is likely that they attribute their behavior to the reward not their intrinsic interest in the activity as the cause of their behavior. This misattribution may lead them to report postbehavior intrinsic motivation that is lower than would be if they had not received the reward.

The attributional theory was expanded by Lepper et al. (1982) to increase the explanatory power of the theory. The undermining effect of the reward increases as the instrumentality between an activity and the reward becomes more salient (Lepper, 1981). If the reward provides positive information about the person's competence, the information may offset part of the instrumentality effect of the reward, leading to a less detrimental effect on intrinsic motivation.

CET and overjustification effect are highly similar in their predictions regarding the undermining effect; even though, they differ in the theoretical processes that they provide. Both theories predict that salient contingent rewards undermine intrinsic motivation. However, overjustification approach focuses only on the instrumentality of the reward as the only factor affecting intrinsic motivation whereas CET focuses on the controlling aspect of the reward as well as its instrumentality. In addition, both approaches predict that the types of rewards that convey positive competence information are less likely to undermine intrinsic motivation. However, CET gives a more complete account of the competence aspect of rewards.

2.5.2.3 Theory of Learned Helplessness

Eisenberger and Cameron (1996) used Seligman's (1975) learned helplessness theory to interpret the undermining effect. This theory suggests that when people receive rewards that are not dependent on their performance (e.g., task-noncontingent or engagement-contingent rewards) they feel helpless because they learn that they cannot control the receipt of the rewards. Many researchers misinterpreted the helplessness concept as the decrements in intrinsic motivation (Deci et al., 1999; Eisenberger & Cameron, 1996). On the other hand, when people receive rewards that are dependent on their performance (e.g., performance-contingent and completion-contingent rewards), they become more industrious. In other words, the reward facilitates learned

industriousness which means that people exert more effort because effortful activity acts as a secondary reward leading to the achievement of more external rewards in the end. Researchers have interpreted this effect as increased intrinsic motivation (Deci et al., 1999; Eisenberger & Cameron, 1996).

The Eisenberger and Cameron's (1996) learned helplessness theory predictions are in sharp contrast with the predictions of CET theory regarding the effect of different types of rewards on intrinsic motivation. It explicitly predicts that engagement-contingent and task-noncontingent rewards undermine intrinsic motivation, whereas performance-contingent and completion-contingent rewards leave unchanged or enhance intrinsic motivation (Deci et al., 1999).

2.5.3 Other Related Theories of Motivation

2.5.3.1 Theory of Fantasy Realization

Based on goal setting theory, contingent rewards can improve effort and performance through several processes including: 1. forcing people to set goals that they otherwise would not, 2. causing people to set more challenging goals, and 3. inducing higher levels of goal commitment in comparison with non-contingent or no reward systems (Locke et al., 1981).

Even though monetary rewards can help people to set more challenging goals that are desirable, they do not guarantee that one actually commits to those goals and strives for their realization. The transition from fantasies and wishes to goal commitment is captured by the theory of fantasy realization. This theory specifies three related self-regulatory processes that people use in order to transform their wishes to actual goal commitments. These processes include mental contrasting, indulging, and dwelling (Fiske et al., 2010; Oettingen, 2000). In

mental contrasting, the individuals first imagine the fulfillment of a fantasy or a desired outcome. Then, they attempt to find ways to move from the present reality to the desired future. Therefore, mental contrasting is a problem-solving strategy. It makes people realize that they have not reached their desires yet and they should take actions to attain their future goals. This will activate expectations of achieving the desired future which in turn, determines the level of goal commitment and subsequent effort for goal pursuit based on the expectations of success. When people have high expectations of success, they actively commit to their goals by exerting a lot of effort to achieve the desired outcomes. However, when expectations of success are low, people refrain from the goals and try to set feasible wishes and desires, so not only mental contrasting facilitates problem-solving strategies to obtain desired goals. But, it also helps people to discriminate between feasible and unfeasible goals and strive toward the feasible ones (Fiske et al., 2010).

There are two other possible routes to goal setting including indulging in positive fantasies and dwelling. Indulging refers to envisioning only the attainment of the desired outcome. Dwelling, on the other hand, refers to reflecting only on the present negative reality. Indulging in positive fantasies involves spending time on visualizing positive mental images of future desired outcomes (Kappes & Oettingen, 2011; Oettingen & Mayer, 2002; Klinger, 1996, 1990). These fantasies depict an idealized situation which includes the attainment of desired outcomes through ideal and smooth processes of achieving these outcomes. Unlike mental contrasting, in indulging strategy, people barely question the feasibility of the desired goal, nor do they consider the possible obstacles, setbacks, or the required effort in the way to obtain those outcomes (Oettingen & Mayer, 2002). In both of these mental strategies the individual fails to

recognize the required actions to achieve the desired outcome. Therefore, they do not activate the expectations of achieving the desired outcome and goal setting does not eventuate in goal commitment and increased effort to obtain the desired future.

Many empirical studies support the predictions of fantasy realization theory (Oettingen et al., 2001; Oettingen, 2000; Oettingen et al., 2000). In one study, Oettingen et al. (2001) tested the effect of the three mentioned strategies on performance. They found that only participants in the mental contrasting condition felt energized, exerted effort and earned higher performance outcomes. Participants in the indulging and dwelling conditions felt moderately energized, invested the lowest level of effort, and received the lowest performance outcomes. In another study, Kappes and Oettingen (2011) found that positive fantasies produce poor performance results because they generate less energy than fantasies that question the desired outcomes.

Based on the fantasy realization conceptualization, the effect of monetary rewards on performance is not completely clear and depends on the mental strategy that the individual selects in order to achieve the desired outcome of high performance and its consequent gain of monetary rewards. Contingent monetary rewards can result in either mental contrasting or indulging strategies. If the individual chooses a mental contrasting strategy by reflecting on the present reality that stands in the way of achieving the desired outcome, expectations of achieving the reward will be activated. This activation will lead to a higher level of goal commitment and subsequent effort to achieve the desired outcome of gaining rewards. However, if the individual gets involved in the indulging strategy, expectations of success will not get activated and the motivational energy towards attaining the desired goal will diminish. The positive fantasy indulgence will make the person focus too much on the reward to the exclusion of possible

strategies that assist in achieving the goal and receiving the contingent reward (Steel & McDonnell, 2012).

In comparison to non-contingent rewards, contingent ones are more likely to generate mental contrasting strategies because the person's level of reward directly depends on their level of performance. Non-contingent rewards, on the other hand, are more likely to result in the indulging and fantasizing strategies since the reward is not related to the performance outcomes.

2.5.3.2 Resource Allocation Theory

Resource allocation theory was developed by Kanfer and Ackerman in 1996 (Kanfer & Ackerman, 1996). Based on this theory, a human being has a limited amount of cognitive capacity. It is important how to allocate this limited capacity because one can either focus on the outcome or on performing the task that will lead to the desired outcome. Misallocation of cognitive resources to the reward rather than the process of achieving it can misdirect motivational resources away from focusing on the task and prevent from attaining the desired reward. This effect can become worse when the level of reward increases. In one study done by Chib et al. (2012), the researchers found that performance increased initially with the presence of reward but it declined as the level of reward level became very high.

A high level of monetary rewards can lead to the "choking under pressure" phenomenon as well (Baumeister, 1984). In this situation, high level of effort as a result of high reward levels will result in a decrement in performance. There are several psychological mechanisms that could lead to choking under pressure including the misallocation of cognitive resources and preoccupation with the reward, supra-optimal levels of arousal (Yerkes & Dodson, 1908),

shifting mental processes from automatic to controlled (Camerer et al., 2005; Langer & Imber, 1979), and narrowing focus of attention (Easterbrook, 1959).

2.5.3.3 Construal Level Theory

Based on construal level theory and affective forecasting, it is difficult to predict what is desirable to an individual ahead of time. The theory suggests that a human being uses different types of information to make decisions for the distant future in comparison with the present or the near future (Trope & Liberman, 2003). When making decisions for the distant future, people focus on their desires and use their goals to make choices. However, as they move towards the goals they have set and get closer to them, they realize the reality clearly and may find that they were not actually interested in the goals (Steel & McDonnell, 2012; Liberman et al., 2007).

In addition to construal level theory, research in affective neuroscience explains the wide gap between wanting and liking. Wanting, liking and learning parts of rewards are processed in distinct neuroanatomical and neurochemical brain reward systems (Berridge, Robinson, & Aldridge, 2009). Thus, it is possible for individuals to pursue rewards that they actually do not like or like the rewards that they never wanted (Steel & McDonnell, 2012). As Kent Beridge stated, "It is relatively hard for a brain to generate pleasure, because it needs to activate different opioid sites together to make you like something more. It’s easier to activate desire, because a brain has several ‘wanting’ pathways available for the task. Sometimes a brain will like the rewards it wants. But other times it just wants them” (University of Michigan News Service, 2007).

2.5.3.4 Social Cognitive Theory

Based on social cognitive theory, an individual's behavior is regulated by an interplay of self-efficacy and outcome expectation. Self-efficacy refers to the "judgment of one's ability to organize and execute given types of performances" (Bandura, 1997). Outcome expectations, on the other hand, refers to the "judgment consequence such performances will produce". In other words, based on social cognitive theory, an individual's performance is determined by both personal cognition, i.e., self-efficacy and outcome expectations such as reward systems.

According to SCT, self-efficacy is the most central mechanism of human regulation. People's beliefs in their efficacy influence their choices, their aspirations, the amount of effort they exert to achieve their goals, the level of their perseverance in the face of difficulties, their thought patterns as self-hindering or self-aiding (Bandura, 1991a).

Self-efficacy can also affect the goal-setting process of self-regulation. The persons' level of perceived competence and self-efficacy positively affects their level of goal setting and goal commitment. The higher the person's level of self-efficacy, the higher their goals will be and the more firmly committed they will be to these set goals (Bandura, 1991b; Locke & Latham, 1990). In addition to its effect on goal setting and goal commitment, perceived self-efficacy would affect the valuation of activities in a way that people display more interest in activities based on the judgment of their self-efficacy in those activities (Bandura & Schunk, 1981).

2.6 Conclusion

Chapter 1 presented an overview of the motivation literature and the related debate on the effect of external reward on intrinsic motivation and performance. The next chapter focuses on

our first paper on the effect of external reward and intrinsic motivation on individuals' overall motivation and performance on an assigned task.

First, we will discuss the limitations of the previous studies on external reward and intrinsic motivation literature. Then, we design a new experiment to address some of the limitations of the past studies. We will see how intrinsic motivation and external reward affect overall task motivation and performance.

Chapter Three: Experimental Study on External Reward, Intrinsic Motivation, Overall Task Motivation, and Performance

3.1 Introduction

Despite the large body of literature in both psychology and economics that is concerned with the potential effects of external rewards and intrinsic motivation on an individual's overall motivation toward and performance in a task, there is still an unresolved debate about the effect of performance-contingent monetary rewards − for which the level of performance determines the amount of the reward − on motivation and performance.

Performance-contingent rewards are the most interesting type of rewards because they can influence motivation and performance in two opposing ways: They can have the strongest negative impact on performance in comparison with other tangible rewards due to their controlling nature, yet they can have a substantial positive effect on performance by conveying competence information when the person performs well in the task (Deci et al., 2001). In this study, we will examine the effect of performance-contingent monetary rewards and intrinsic motivation on overall motivation and performance in an experimental setting. Overall motivation, as defined in this study, refers to the sum of extrinsic motivation and intrinsic motivation that produces a total motivation to accomplish a task.

In this study, we examine the effect of performance-contingent monetary reward and intrinsic motivation on overall task motivation and performance in an experimental setting. There are two main, and contrasting, streams of research in the literature. The first stream of research, conducted by cognitive evaluation theorists in psychology, focuses on the effect of different types of external rewards on an individual's intrinsic motivation (for a review of these

studies, refer to Deci et al., 1999). The results of this stream show that external rewards, especially in the form of performance-contingent rewards, produce an "undermining effect" (Deci & Ryan, 1980) on intrinsic motivation, which was implicitly or explicitly interpreted by many authors as decrements in task motivation and performance (Gneezy & Rustichini, 2000).

On the other hand, the second stream of research, which has mostly been conducted by behavioral psychologists and economists, is concerned mainly with the effect of external rewards on performance (Camerer & Hogarth, 1999). These studies indicate that performance-contingent rewards improve performance. The positive effect of performance-contingent rewards on performance as found in these studies was extended to the conclusion that external rewards improve motivation, through which performance improves.

Although these two streams of research seem to produce contradictory results regarding the effect of performance-contingent rewards on motivation and performance, we will show that there are several fundamental distinctions between them that make it impossible to compare their results. In fact, these two streams are not really intersecting. Studies in the first stream mostly focus on the effect of external reward on intrinsic motivation, while studies in the second stream focus on the effect of external reward on performance. Previous studies in both streams have not simultaneously examined the effect of performance-contingent rewards and intrinsic motivation on overall task motivation and performance.

Furthermore, intrinsic motivation in past studies has been measured through either self-reports of interest or free-choice behavior. Free-choice behavior refers to the degree to which participants engage in the target task during a free-choice period subsequent to the experimental phase (Deci, 1971). In most studies, the free-choice behavior was assessed based on the amount

of time spent on the target task during the free-time period (e.g. Amabile et al., 1986). Self-reports of interest, on the other hand, consist of questions assessing interest and enjoyment in the experimental activity. Each of these measures has methodological deficiencies, which we discuss below.

In order to resolve these difficulties, we designed and ran an experiment to examine the effect of performance-contingent monetary rewards and intrinsic motivation on overall motivation and performance. By making use of participants' self-selection into an area of knowledge, hypothesizing that those who self-select themselves into/out of a specific area of knowledge have high/low intrinsic motivation towards that area, and matching/mismatching participants' fields of study with the topic of their test, we avoid the confounding effects seen in previous studies and obtain results to test the effect of external reward and intrinsic motivation on overall task motivation and performance.

3.2 Past Studies' Limitations

Experimental studies conducted within both the additive and undermining effect perspectives have some limitations. First of all, most of the previous experimental studies have either examined the effect of external reward on intrinsic motivation or the effect of external reward on performance. Only a few studies have examined the simultaneous effect of external reward and intrinsic motivation on performance, and none has specifically examined the effect of performance-contingent monetary rewards in particular.

The first stream of experimental studies, which is in line with standard economic theories and behavioral theories, has mostly attempted to examine the effect of external reward on effort and performance (for a review of these studies, please refer to Bonner et al., 2000; Camerer &

Hogarth, 1999). The independent variable in these studies is contingent reward and the dependent variable is performance. Intrinsic motivation has neither been included nor operationalized in these studies. Change in the level of intrinsic motivation is inferred from change in the level of performance (Ariely et al., 2009; Gneezy & Rustichini, 2000). Also, effort is not separately measured and studied in these experiments and is treated as the operationalization and manifestation of the individual's motivation (Ariely et al., 2009).

These experiments showed that a sufficiently high level of performance-contingent monetary incentive increases performance (Frey & Jegen, 2001; Gneezy & Rustichini, 2000; Lazear, 2000). Other experiments have also shown that if the payment level is acceptable, it increases performance (Lazear, 2000). But, paying a small amount that is less than the amount expected by the participant for the completed task may change the perceived nature of the task and result in a decline in motivation and performance (Heyman & Ariely, 2004; Frey & Jegen, 2001; Gneezy & Rustichini, 2000). On the other hand, paying an excessively high level of monetary incentive can negatively impact performance due to the "choking under pressure" phenomenon (Ariely et al., 2009).

As an example of experimental studies in this stream, the study done by Gneezy and Rustichini (2000) can be mentioned. Using a laboratory experiment, the researchers examined the effect of monetary rewards on performance. The findings showed that the effect of monetary rewards on performance was not monotonic. When the level of payment was sufficient, performance significantly increased relative to the no-payment or low-payment treatments. However, when the level of payment was too low, it decreased performance in comparison with

the no-payment condition. On the other hand, increasing the amount of rewards did not have a significant effect on performance after some payment level.

Similar to Gneezy and Rustichini's (2000) study, Ariely et al.'s study (2009) examined the effect of different levels of contingent financial rewards on performance. In addition, they examined the generality of any detrimental effect of monetary rewards on performance by using different types of tasks in their experiment. The tasks were primarily based on motor skills, memory, creativity, and physical effort. The results of this experiment showed that very high levels of performance-contingent monetary rewards produced lower performance on the first three types of tasks due to mechanisms leading to choking under pressure. However, for the tasks that only required physical effort, higher level of performance-contingent monetary rewards improved performance (Ariely et al., 2009).

The second stream of experimental studies, both in psychology and economics, examines the predictions of CET and other related theories regarding the undermining effect of external reward on intrinsic motivation. Financial reward is the independent variable and intrinsic motivation is the dependent variable in these studies.

Subsequent to some preliminary studies conducted by Deci (1972 a, 1972b, 1971), a large number of laboratory experiments in psychology concentrated on testing the effect of external reward on intrinsic motivation; several meta-analytical studies have been conducted to examine the undermining hypothesis (Tang & Hall, 1995; Wiersma, 1992; Rummel & Feinberg, 1988). However, Deci et al.'s (1999) meta-analysis showed that some of these early studies were flawed from a methodological perspective. The existence of methodological issues in these studies

resulted in an increase in the acceptance of behaviorist approach including the meta-analytical studies of Cameron and Pierce (1994) and Eisenberger and Cameron (1996).

Deci et al.'s (1999) meta-analysis is the best available study on the effects of external rewards on intrinsic motivation in social psychology. It reviews a total of 128 different experiments from 1971 to 1997, including all the studies considered in the meta-analyses done by Cameron, Pierce, and Eisenberger. The results of the meta-analysis generally supported the main predictions of CET regarding the undermining effect of contingent rewards on intrinsic motivation. Among different types of rewards, CET specifically asserts that performance-contingent monetary rewards will undermine intrinsic motivation by making the controlling aspect of the external reward salient (Deci, 1972a, 1972b, 1971). The controlling aspect undermines feelings of self-determination and decreases intrinsic motivation. However, these studies have several limitations that require further examination.

First of all, performance is not reported as a dependent variable in most of these studies. For example, in none of Deci's experimental studies are the performance results reported for the experimental task (Calder & Staw, 1975a; Salancik, 1975). Thus, it is unclear whether any increase or decrease in the level of intrinsic motivation has actually improved or undermined performance.

One of the underpinning assumptions of CET is that there is a positive relationship between intrinsic motivation and performance (Pinder, 1975). However, only a few studies have attempted to examine the real effect of intrinsic motivation on performance (e.g., Daniel & Esser, 1980; Hamner & Foster, 1975; Calder & Staw, 1975b). Examining the effect of external reward

on intrinsic motivation without measuring performance gives a partial view of the relationship between motivation and performance (Calder & Staw, 1975a; Pinder, 1975; Salancik, 1975).

Second, experimental studies that have examined the effect of external reward on intrinsic motivation have used two main measures for intrinsic motivation: self-reports of interest and free-choice behavior. These measures have several disadvantages.

The most important disadvantage of self-reports of interest is their subjectivity. Self-reports of interest may be biased due to other-regarding preferences. Since participants know that the experimenter will see and analyze their answers on the self-report of interest and since the questions are transparent, participants may try to give answers that satisfy the experimenter rather than their true feelings, so demand characteristics and interpersonal factors can influence self-report measures of intrinsic motivation (Deci et al., 1999). In addition, since participants complete self-reports of interest after the task, they are aware of the type of external reward at the time of writing the report, so the use of self-reports of interest for a rewarded task may result in a confusion of the enjoyment of external reward with the individual's interest in the task (Deci et al., 1999), so that these measures are confounded.

In addition, studies using self-reports of interest do not report any pretreatment measure of intrinsic interest. In a typical experiment, participants take part in a task assumed to be intrinsically motivating by the researcher—the tasks used in these studies included: solving SOMA puzzles, the most widely used task in the literature (Williams, 1980; Enzle & Ross, 1977; Deci, 1976; 1972a; 1971); drawing pictures with colored markers (Lepper et al., 1973); building with erector sets (Kruglanski et al., 1971; Wimperis & Farr, 1979; Pinder, 1976); and playing word games (Weiner, 1980; Cameron & Pierce, 1994; Rummel & Feinberg, 1988).

Another methodological issue of self-reports is the lack of reliability. Most studies have used only one item (or very few items) in their measure of intrinsic motivation; this reduces the reliability of such measures (Deci et al., 1999).

With free-choice behavior, there is a possibility that participants take part in the task only because there is no alternative for them to do another task or because their reward is dependent on the completion of the task (Gneezy & Rustichini, 2000). In such cases, participants' free-choice behavior cannot be used as a measure or an indicator of their intrinsic motivation.

Another important problem is that performance is not reported in free-choice studies of intrinsic motivation as in studies that used self-reports of interest, so it is not clear that the difference in free-choice behavior is due to participants' differences in intrinsic motivation or differences in performance (Calder & Staw, 1975a). One argument would be that the presence of monetary reward during the main task may encourage participants to increase their effort and performance during the experimental session. Therefore, decrease in the amount of time spent on the task in the free time period—free-choice behavior—could be attributed to factors such as satiation or fatigue rather than any changes in the level of intrinsic motivation (Calder & Staw, 1975a).

In some experimental settings, free-choice behavior may be influenced by external reward. This may happen when the experimenter gives the external reward on condition that the participant completes the task. In this case, the participant is convinced to do the task and his/her performance is influenced by the reward he/she will receive after task completion.

Although the main stream of studies both in behavioral and cognitive evaluation theories has focused on only part of the relationship between different motivational factors and

performance, a few researchers have attempted to address these issues by testing the interaction effect of external reward and intrinsic motivation as independent variables on related dependent variables. We will briefly review some of these studies here and try to clarify some methodological difficulties regarding these studies.

In one study, Calder and Staw (1975b) designed an experiment to test the interaction between completion-contingent external reward and intrinsic motivation by manipulating both intrinsic motivation and external reward as two independent variables. In this experimental design, intrinsic motivation and external reward were clearly manipulated, so the design gave the ability to test the additivity versus the interaction hypothesis between intrinsic motivation and external reward (Calder & Staw, 1975b).

The experimental task included solving 15 jigsaw puzzles of either boring or interesting type. To manipulate intrinsic motivation, the researchers used the concept of boring versus interesting activity. They used blank puzzles in one group (low intrinsic motivation) and puzzles with interesting pictures in another group (high intrinsic motivation). External reward had two categories of no pay versus the completion-contingent payment of $1 for completing the experimental task. Participants in the payment group received their $1 payment after completing all the 15 jigsaw puzzles. The dependent variables of this study were task satisfaction and task persistence.

This study is one of the few studies that have attempted to manipulate intrinsic motivation (Deci et al., 1999). But, it has some methodological issues. First, the type of task used in this study, as in many studies on intrinsic motivation, is solving puzzles. Puzzle solving is not a good representation of real world tasks, especially if the participant is not familiar with puzzle-

solving prior to the experiment (Hamner & Foster, 1975). Second, the manipulation check for intrinsic motivation is done after the task is complete without any pretreatment measure for intrinsic motivation (Cameron & Pierce, 1994). The study uses after-the-task self-reports of interest in order to check the manipulation for the independent variable of intrinsic motivation. This type of manipulation check mixes the effect of reward and intrinsic motivation. Thus, one cannot find out whether the manipulation for intrinsic motivation has actually worked out. Finally, this study does not measure participants' performance. As a result, the effect of external reward and intrinsic motivation on performance is still missing in this study.

In another study, Hamner and Foster (1975) designed experiments to examine the interaction effect of engagement-contingent and completion-contingent rewards and intrinsic motivation. Similar to Calder and Staw (1975b), Hamner and Foster (1975) attempted to clearly manipulate both external reward and intrinsic motivation as independent variables of the study. The study has a 2x3 design with two levels of task interest (boring versus interesting task) and three levels of payment (no-pay, completion-contingent pay and engagement-contingent pay).

The boring task was a task where the subjects coded and transferred the scores from a math survey (with fake responses) of college females to a FORTRAN worksheet. The interesting task was a task where the subjects coded and transferred the scores from a sexual attitude survey (with fake responses) of college females to a FORTRAN work sheet.

In the no-pay condition, participants were not paid for completing this task and acted as the control group in this experiment. In the engagement-contingent condition, participants received 75¢ for participating in the task for 20 minutes. In the completion-contingent condition, subjects received 5¢ for scoring each respondent questionnaire.

The study examines the effect of the interaction between external reward and intrinsic motivation on task motivation and performance. There were four focal dependent variables in this study: the level of performance output (number of items scored), quality of output (number of items scored incorrectly), interest in the task on a scale ranging from extremely boring to extremely interesting, and satisfaction with the payment. In comparison with Calder and Staw's study, this study uses a kind of task that is closer to the real-world tasks. In addition, the study measures performance in different treatment conditions which gives the ability to compare the effect of external reward and intrinsic motivation on task motivation and performance.

However, like Calder and Staw's (1975b) study, the manipulation check for intrinsic motivation is done after at the experimental session. It is done using a single-item measure for the level of interest in the task in the post-experimental questionnaire. Participants filled the post-experimental questionnaire after they completed the task and received their payment. Thus, interest in the task (level of intrinsic motivation) is mixed with the enjoyment of reward and the participant's level of performance.

Similar to the above-mentioned studies, other studies have also attempted to test the interaction effect of external reward and intrinsic motivation as two manipulated independent variables on related dependent variables (Pinder, 1976). These studies all have methodological problems such as the ones mentioned in Calder and Staw (1975b) and Hamner and Foster (1975) studies. Since self-reports of interest are completed after the experiment, it is possible that the enjoyment of monetary reward be confounded with the interest in the task in the reward treatments. As well, some studies use only job satisfaction and overall motivation as the

dependent variables. Lack of information about participants' performance in these studies makes it impossible to examine the effect of external reward and intrinsic motivation on performance.

3.3 Hypotheses

Considering the limitations of the previous studies on external reward, intrinsic motivation and performance, our research examines the effect of external reward x intrinsic motivation interaction on performance using self-selection into a specific area of knowledge as a measure for assessing intrinsic motivation.

3.3.1 Self-Selection and Intrinsic Motivation

In this experiment, we developed a new measure for assessing intrinsic motivation by matching/mismatching participants' field of study in a university setting with the topic of a test in which they will take part. We presume that individuals have self-selected into their field of study based, at least in part, on their intrinsic motivation to learn more about that specific field of knowledge. We hypothesized that the participants whose field of study matches (mismatches) with the topic of their test would have high (low) intrinsic motivation towards the test. While other factors including native and personal ability, need for less effort, job availability, and social pressures may also affect an individual's self-selection (Zhang, 2007), we presume that intrinsic interest will be highly correlated with self-selection into an area of knowledge.

Research on major selection has found that students' selection of different majors is influenced by the same set of factors (Lee & Lee, 2006; Kim et al., 2002). These factors can be divided into two categories: experiential and instrumental factors. While instrumental factors relate to the costs and benefits of choosing a major, experiential factors relate to the beliefs regarding future pleasure and satisfaction stemming from this major choice. In other words,

experiential beliefs are related to the expectation of enjoyment that may result from choosing a specific major. Based on previous research, a genuine interest in a field is one of the most important factors influencing students' major selection (Zhang, 2007; Malgwi et al., 2005; Adams et al., 1994). Thus, our first two hypotheses relate to major selection and its effect on intrinsic motivation:

Hypothesis 1a. Individuals whose field of study matches the type of the test in which they will take part have high intrinsic motivation towards the test (high intrinsic motivation group).

Hypothesis 1b. Individuals whose field of study mismatches the type of the test in which they will take part have low intrinsic motivation towards the test (low intrinsic motivation group).

This new measure for intrinsic motivation has several advantages in comparison with previous measures of intrinsic motivation. First, our measure is based on a more realistic task type than the use of interesting/boring puzzles in previous studies of intrinsic motivation. Participants take part in a task which is familiar to them and is based to a greater or lesser degree on their proven real interests over the long run. Second, this measure is objective. It is not based on the subjective reports of participants. Third, the check to see how the field of study indicated intrinsic interest in the type of test is done at the beginning, not the end. This type of manipulation check will solve the main methodological issue of the previous studies on intrinsic motivation, which were based on self-reports of interest for the measurement or manipulation check of intrinsic motivation. Therefore, our measure is not confounded with monetary rewards.

Using this new measure for intrinsic motivation, we will examine the effect of intrinsic motivation, external reward, and their interaction on performance.

3.3.2 Performance-Contingent Reward and Performance

As explained in the literature review section, most theories of motivation including reinforcement theory, expectancy theory and agency theory assume a positive effect of performance-contingent financial reward on performance. Experimental studies have supported these theories by finding that a sufficiently high level of external reward improves performance (Frey & Jegen, 2001; Gneezy & Rustichini, 2000; Lazear, 2000). As an example, Gneezy and Rustichini (2000)'s experiment showed that when the level of external reward was sufficient, performance significantly increased relative to the no-payment or low-payment treatments. Other experiments have also shown that if the payment level is acceptable, it can increase performance (Lazear, 2000). Based on the above-mentioned theories and the related experiments, we hypothesize that performance-contingent monetary rewards will improve performance.

Hypothesis 2. Performance-contingent monetary rewards improve performance.

3.3.3 Intrinsic Motivation and Performance

Although there is a large body of literature on the effect of external reward on performance, the relationship between intrinsic motivation and performance has not been clearly addressed by researchers. The literature on intrinsic motivation is mainly focused on the effect of external reward on intrinsic motivation. This is due to the implied assumption made by cognitive psychologists that intrinsic motivation is an inherent interest towards task accomplishment and as a result, it will automatically improve performance (Ryan & Deci, 2000). However, only very few experimental studies have examined this relationship (Pinder, 1976; Hamner & Foster,

1975). In this study, we examine the relationship between intrinsic motivation and performance. Based on the cognitive evaluation theory argument, we hypothesize that intrinsic motivation improves performance.

Hypothesis 3. Intrinsic motivation improves performance.

3.3.4 Intrinsic Motivation x External Reward Interaction and Performance

In line with additive theories of motivation such as reinforcement theory, expectancy theory, agency theory and goal setting theory, we hypothesize that effects of intrinsic motivation and external reward are additive, and their interaction have a synergistic effect on performance. In other words, individuals who have high intrinsic motivation will significantly improve their performance in the presence of external reward due to a higher level of total motivation. Figure 3.1 illustrates the first four hypotheses on the effect of external reward, intrinsic motivation, and their interaction on performance.

Hypothesis 4. Individuals with high levels of intrinsic motivation increase their performance more in the presence of external reward in comparison with those with low levels of intrinsic motivation.

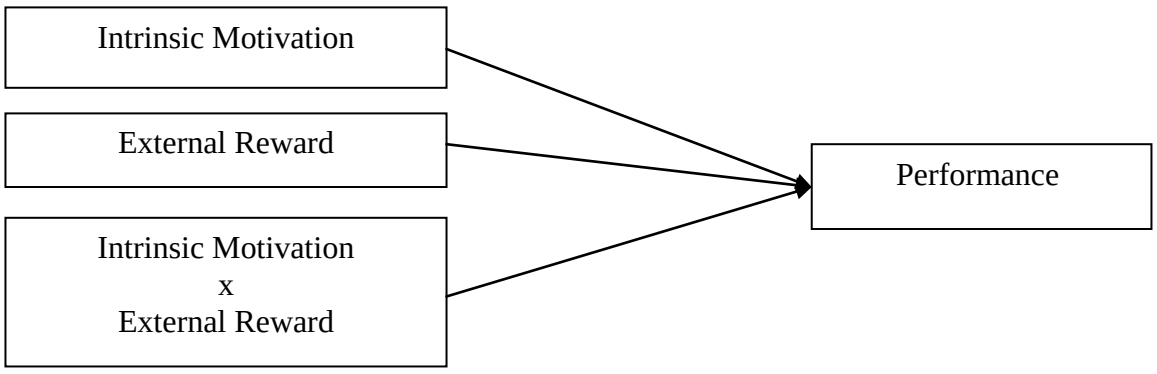


Figure 3.1 Intrinsic motivation, external reward, and performance

3.3.5 Intrinsic Motivation and Task Motivation

Based on cognitive evaluation theory, intrinsic motivation is highly correlated with overall task motivation. Therefore, we hypothesize that intrinsic motivation improves overall task motivation.

Hypothesis 5. Intrinsic motivation improves overall task motivation.

3.3.6 External Reward and Task Motivation

There is a debate regarding the effect of performance-contingent monetary rewards on overall task motivation. As explained in the theoretical chapter of our paper, most behavioral theories of worker motivation including expectancy theory (Vroom, 1964), reinforcement theory (Bindra, 1959; Skinner, 1953), and agency theory (Eisenhardt, 1989) propose that monetary rewards have a positive effect on motivation. However, theories such as cognitive evaluation theory (Deci et al., 1999) and the theory of learned helplessness (Eisenberger & Cameron, 1996; Seligman, 1975) argue that performance-contingent monetary rewards decrease motivation due to their highly controlling nature. Based on cognitive evaluation theory, performance-contingent rewards have a high tendency to undermine motivation because the reward is linked to performance and as a result, it can strongly control people's behavior (Deci et al., 1999). Considering these two arguments, we developed two competing hypotheses:

Hypothesis 6a. Performance-contingent monetary rewards improve overall task motivation.

Hypothesis 6b. Performance-contingent monetary rewards undermine overall task motivation.

Figure 3.2 illustrates the first four hypotheses on the effect of external reward, intrinsic motivation, and their interaction on task motivation. Table 3.1 provides a list of all the variables that were measured and included in this experiment.

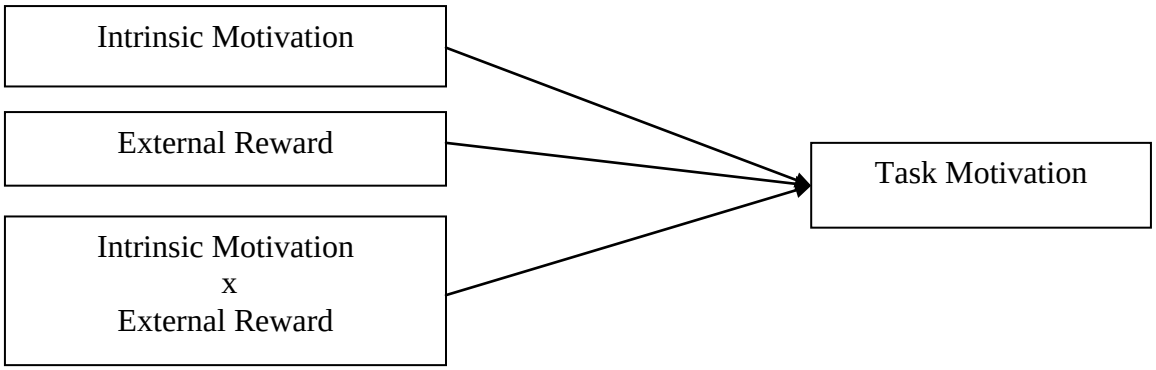


Figure 3.2 Intrinsic motivation, external reward, and task motivation

Table 3.1 Measured variables

Variable Name	Variable	Measure	Variable Type
Performance	Dependent	Number of correct answers	Scale (0-30)
Overall Motivation (BehMotiv)	Dependent	Amount of time spent on the task	Scale (0-50)
Mood	Dependent	Two items selected from PANAS (Positive and Negative Affect Scale) mood scale	Scale
Intrinsic Motivation (IntMotiv)	Independent	Match/mismatch between field of study and test type	Nominal with two levels (contrast coding): -0.5 = Low; 0.5 = High
Intrinsic Motivation Manipulation Check	Dependent	Average score on the intrinsic motivation scale	Scale (1-7)
External Reward (ExtRew)	Independent	No monetary incentive/monetary incentive	Ordinal with two levels (contrast coding): -.05 = Fixed pay ; 0.5 = Piece rate
IntMotivxExtRew	Independent	Interaction between intrinsic motivation and external reward	Ordinal with two levels (contrast coding) Reference: FixedLow and PieceHigh: 0.25 FixedHigh and PieceLow -0.25
Knowledge	Covariate	Number of math and English courses completed in the university	Scale
Confidence	Covariate	Confidence level	Scale (1-7)
PureGuess	Covariate	Percentage of answers based on pure guessing	Scale (0-100%)
Age	Covariate	Participant's age	Scale
Gender	Covariate	Participant's gender	Nominal with two levels (dummy coding): 0 = Female; 1 = Male
Field	Covariate	Participant's field of study	Nominal with two levels: 0 = English-related; 1 = Math-related
Under/Graduate (UnderGrad)	Covariate	Whether the student is undergraduate or graduate	Nominal with two levels: Undergraduate = 1; Graduate = 0
Primary Language (EnglishFirst)	Covariate	Participant's primary language	Nominal with two levels (dummy coding): 1 = English; 0 = Other
Ethnicity	Covariate	Participant's ethnicity	Nominal with 7 categories (dummy coding): Caucasian, First Nations, Biracial, Black, Asian, Middle Eastern, and Other

3.4 Experimental Design

We would like to examine the effect of external reward and intrinsic motivation on performance and also to determine whether self-selection into an area of knowledge provides an objective measure for intrinsic motivation.

To test these hypotheses, we performed an experiment using a 2 (external reward type: fixed payment vs. piece-rate payment) x 2 (intrinsic motivation type: match vs. mismatch) x 2 (field: math-related vs. literature-related) experimental design. In the next sections, we first describe the manipulation of intrinsic motivation. We then define variables included in the study and explain the procedure of the study and the study results.

3.5 Manipulating Intrinsic Motivation

In order to manipulate the level of intrinsic motivation toward a given task, we used self-selection into a specific area of knowledge. Students from mathematics related departments—mathematics and statistics, electrical engineering, mechanical engineering, chemical and petroleum engineering, computer engineering and computer science, geomatics engineering and civil engineering—and literature-related departments—art, dance, and drama; English; French, Italian, and Spanish; Germanic, Slavic, and East Asian studies; Greek and Roman Studies; education; history; philosophy; and archeology—were selected and randomly assigned to two different tests: mathematics and English language. High intrinsic motivation treatments were comprised of students in math-related/literature-related majors who were randomly assigned to take mathematics/English test (match treatments). Low intrinsic motivation treatments, on the other hand, were comprised of students in math-related/literature-related majors who were randomly assigned to take a test of English language/mathematics (mismatch treatments). We

hypothesized that students in the match treatments would have high levels of intrinsic motivation, while students in the mismatch treatments will have low levels of intrinsic motivation.

3.6 Variables

3.6.1 Main Dependent Variables

The study had two main dependent variables: 1. performance and 2. overall motivation. Performance was measured by the number of correctly answered questions in a mathematics or English language test. The mathematics/English test consisted of 30 multiple-choice questions with one definite correct answer for each question. Students were given 50 minutes to answer these questions. The number of correct answers in the mathematics/English test was used as a measure for participants' performance. The number of correct answers was calculated after the completion of the experiment, and students in the piece-rate payment treatment received a performance-contingent reward based on their performance ($0.50 per each correctly answered question).

Overall motivation was measured by use of a behavioral measure for motivation: the amount of time spent on the test of math/English language. We used a behavioral measure since it is a more reliable measure for assessing motivation due to its unobtrusiveness (Deci et al., 1999). Unlike self-report of interest, it is unobtrusive; participants do not believe that the experimenter has collected data regarding the amount of time they spent on the task (Deci et al., 1999), so interpersonal factors will have a minimal impact on this measure of motivation (Deci et al., 1999).

3.6.2 Independent Variables

The study has two focal independent variables: intrinsic motivation and external reward. These variables are operationalized as binary dummy variables. Intrinsic motivation has two levels of high and low intrinsic motivation. Participants in the match treatments were hypothesized to have high intrinsic motivation and participants in the mismatch treatments were hypothesized to have low intrinsic motivation. External reward has two levels of fixed or piece-rate. Participants in the piece-rate treatments were paid $.5 per correct answer in addition to their $15 participation fee. Participants in the fixed payment treatments received only $15 as their participation fee.

In order to test our three research hypotheses related to the main effect of intrinsic motivation, main effect of eternal reward and the interaction between intrinsic motivation and external reward, we used a set of contrast codes. Contrast coding is the most appropriate method for these types of research hypotheses (Cohen & Cohen, 1985). We coded participants in the high intrinsic motivation treatment +0.5 and -0.5 for those in the low intrinsic motivation treatment. In addition, we used +0.5 for participants in the piece-rate treatment and -0.5 for participants in the fixed pay treatment. The interaction between intrinsic motivation and external reward is a function of the product of these codes. Contrast codes are useful for interactions of categorical variables because they are orthogonal and they represent meaningful differences between means of different group conditions (Cohen & Cohen, 1985).

Dummy variables are similar to uncentered quantitative variables. The use of dummy variable coding instead of contrast coding may make the results difficult to interpret. The significance of the interaction term is not influenced by the use of dummy variable coding

scheme. However, the absence of centering which is characteristic of contrast coding makes the main effects of the variables difficult to interpret (Cohen & Cohen, 1985).

3.6.3 Additional Control Variables

In addition to the primary measures of interest, we also gathered information on variables that may be affected by our manipulation, to rule out other explanations of our results and any confounding of our independent variables with other variables. First of all, there is a possibility that our manipulation for intrinsic motivation affects individuals' level of confidence. In order to check for this problem, we asked all participants to indicate their confidence level prior to the mathematics/English test using one measure ("Please indicate how confident you are that you will perform well in this test."; 1 = not at all, 7 = very much). Second, there is a possibility that knowledge may be heterogeneous in different treatments. Since our manipulation for intrinsic motivation is based on the match between field of study and test type, participants in different treatments may have heterogeneous levels of knowledge regarding the topic of the test in which they will take part. To solve this problem, we controlled for knowledge by measuring the number of math and English courses that the participant had completed in the university. Data on other factors including age, gender, etc was also collected.

3.6.4 Other Dependent Variables

To test the likelihood of the effect of the test on participants' mood and to eliminate the possibility of other reasons for the observed levels of performance, we measured mood after the test for a sub-sample of 20 participants and checked whether our manipulation for external reward had a significant effect on mood in different treatment groups. Mood was measured by use of two items selected from PANAS scale (Positive and Negative Affect Scale) (Watson et

al., 1988). After the completion of the test, participants indicated how much they agreed with the following two sentences on a 1-7 scale: 1. I felt **happy** while I was taking the test and 2. I felt **sad** while I was taking the test. The results of a Univariate Analysis of Variance (ANOVA) on the effect of intrinsic motivation and external reward on mood showed that our manipulation for external reward (F = 0.387, p = 0.547) and intrinsic motivation (F = 1.923, p = .185) did not have a significant effect on mood.

3.7 Participants and Procedure

In order to test our hypotheses, 92 participants (57 males and 35 females) from a variety of departments in the University of Calgary were selected and randomly assigned to the conditions of a 2 (intrinsic motivation: high vs. low) x 2 (external reward: piece rate vs. fixed payment) x 2 (field: mathematics-related vs. literature-related) between-subjects factorial design. Participants' fields of study were categorized as either math-related or literature-related majors. We hypothesized that participants in the math-related majors would have a high intrinsic motivation towards taking a math test and a low intrinsic motivation towards taking an English test. In addition, participants in the literature-related majors would have a high intrinsic motivation towards taking an English test and a low intrinsic motivation towards taking a math test.

In order to examine whether the manipulation for intrinsic motivation had worked, in the beginning of the experiment, participants were asked two questions regarding their intrinsic interest prior to the English/mathematics test. In addition, they also answered one question regarding their level of confidence. We measured confidence to examine the possibility of the

effect of intrinsic motivation manipulation on confidence among participants in the match vs. mismatch treatments. Appendix A.1 presents the preliminary questionnaire.

After answering these questions, participants took an online test. In this step, participants in the piece-rate treatments were reassured of the amount of monetary reward ($0.50 per correct answer) they would receive for taking part in the test. The test was installed on Blackboard, the university's standard course management system. Blackboard is designed to facilitate communication between the course instructor and students. Using Blackboard gave us the ability to run online tests that were automatically graded by the system immediately after the test was submitted (http://elearn.ucalgary.ca/blackboard). The test included questions in either mathematics or English language. The questions in math and English tests were selected from the Graduate Management Admission Test (GMAT). Participants' performance was measured immediately after the test based on the number of their correctly answered questions in the math or English test. Appendix A.2 provides an overview of the math/English test section of the experiment along with a sample of the mathematics and English questions.

Finally, all participants completed a set of questions regarding their overall motivation, mood, and demographic data. The items related to overall motivation were selected from the Intrinsic Motivation Inventory (IMI) (http://www.selfdeterminationtheory.org/questionnaires/10-questionnaires/50). The Intrinsic Motivation Inventory is a multidimensional measurement instrument that is intended to assess participants' subjective motivation related to a target activity in laboratory experiments. Experiments related to intrinsic motivation and self-regulation frequently use this measure for assessing participants' level of intrinsic motivation towards the experimental task (e.g., Ryan, 1982; Ryan et al., 1983). We used IMI items as a self-report of

participants' overall motivation. We selected some of the IMI's items to assess participants' interest/enjoyment and perceived effort while performing the test. However, we did not use this subjective measure of overall motivation in our study. Instead of that, we used time spent on the task which is a behavioral measure of overall motivation. According to the motivation literature, the behavioral measure is a more valid and accurate measure of intrinsic motivation than the subjective measure of intrinsic motivation, which is based on a self-report of interest in the task (Deci et al., 1999). Demographic questions included information regarding knowledge, age, gender, ethnicity, and primary language. Appendix A.3 provides the after-the-test questionnaire.

After completing the experiment, participants were guided to a separate room to receive their payment as well as feedback regarding their performance in the math/English test they had completed. Participants in the fixed payment treatments received $15 for their participation; the payment was not contingent on their performance in the test. Participants in the piece rate treatments received $0.50 per correct answer in the test in addition to their $15 participation fee.

3.8 Intrinsic Motivation Manipulation Check

To see whether the manipulation for intrinsic motivation as taking part in a test that matched/mismatched one’s area of knowledge worked, in a preliminary questionnaire prior to the main test, participants were asked to indicate the extent to which they were interested in both English language and mathematics. On a 7-point scale ranging from 1 = not at all to 7 = very much, participants answered four interest-related questions including the following questions: (1) Please indicate how interested you are in solving general mathematics problems, (2) Please indicate how interested you are in answering English questions, (3) Please rate how interesting

mathematics is to you in general on scale from 1 to 7, and (4) Please rate how interesting English is to you in general on a scale from 1 to 7.

In order to test the manipulation check statistically, we restructured the data by creating a unique variable called Interest as a substitute for both MathInterest and EnglishInterest. Then, we ran a multiple regression models on the effect of field of study on Interest. We added the related control variables including confidence, knowledge, age, gender, under/grad, and language to the regression model. Table 3.2 shows the results of our multiple regression analysis on the effect of the field of study on math versus English interest. The results show that our manipulation for intrinsic motivation has worked; there was a significant effect of field of study on interest (β = 0.871, p <0.05).

Table 3.2 Intrinsic motivation manipulation check

	Dependent Variable: Interest
	Model
(Constant)	1.241 (0.262)
Field	0.871* 0.205 (0.030)
Confidence	0.233* 0.165 (0.035)
Knowledge	-0.023 -0.51 (0.498)
Age	0.068* 0.167 (0.05)
Gender	-0.225 -0.055 (0.494)
UnderGrad	-0.219 -0.054 (0.507)
EnglishFirst	0.148 0.037 (0.688)
F Statistic	2.746**
P Value	0.010
N	177

Note: We have provided the unstandardized regression coefficients and standardized coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

3.9 Main Statistical Analyses and Results

As the first step, we ran descriptive statistics on all the variables including dependent variables, independent variables, and control variables for all the treatments in total and for each

treatment separately. In total, 45.6% of the participants reported English language as their first language. The majority of the participants were Caucasians including 43.5% of the participants, followed by Asians who comprised 34.7% of participants. 62% of the participants were male and 38% of them were female. There was not a significant difference in the number of males and females in different treatments. Table 3.3 provides descriptive statistics of all the variables in the study for all treatments in total and for each separate treatment.

Table 3.3 Descriptive statistics

Descriptive Statistics	All Treatments			HighIntMotiv-High ExtRew			High IntMotiv-Low ExtRew			Low IntMotive-High ExtRew			Low IntMotive-Low ExtRew		
Variable	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**
Performance	18.71	6.55	92	22.08	4.92	26	20.09	7.33	22	16.41	6.69	22	15.64	5.20	22
BehMotiv	35.95	10.22	92	38.96	11.05	26	34.36	10.02	22	37.27	8.00	22	32.64	10.74	22
IntMotiv	0.022	0.50	92	0.5	0	26	0.5	0	22	-0.5	0	22	-0.5	0	22
ExtRew	0.02	0.50	92	0.5	0	26	-0.5	0	22	0.5	0	22	-0.5	0	22
IntMotivxExtRew	0.01	0.25	92	0.25	0	26	-0.25	0	22	-0.25	0	22	0.25	0	22
Confidence	5.09	1.46	92	5.58	1.10	26	6.05	0.84	22	4.23	1.27	22	4.41	1.71	22
Knowledge	2.67	4.42	91	4.81	6.71	26	4.00	3.35	21	0.86	1.86	22	0.68	0.99	22
PureGuess	24.35	27.08	92	24.31	30.80	26	22.36	28.48	22	29.09	27.76	22	21.64	20.77	22
Age	24	4.86	90	24.62	6.61	26	22.95	3.15	22	23.57	4.30	21	24.76	4.37	21
Gender	0.62	0.49	92	0.54	0.51	26	0.59	0.50	22	0.68	0.48	22	0.68	0.48	22
UnderGrad	0.54	0.50	92	0.5	0.51	26	0.68	0.48	22	0.59	0.50	22	0.41	0.50	22
EnglishFirst	0.46	0.50	92	0.54	0.51	26	0.5	0.51	22	0.45	0.51	22	0.32	0.48	22
Field	0.67	0.47	92	0.58	0.50	26	0.68	0.48	22	0.73	0.46	22	0.73	0.46	22
Mood	3.68	1.36	20	3.90	0.74	7	4.20	1.82	5	2.92	1.64	4	3.42	1.48	4
Valid N (listwise)			19			7			5			4			3

In order to test hypotheses 2 to 4, we ran multiple regression models in a step-by-step format. Table 3.4 illustrates the results of the three multiple regression analyses. In model 1, we only included the control variables. The results of this regression model indicated that the covariate Confidence (β = 1.598, p = 0.001) had a significant positive effect on performance. Another significant covariate was PureGuess, defined as the percentage of answers based on pure guessing (β = -0.089, $p < 0.001$) This result was theoretically sound; as participants answered questions based on pure guessing instead of answering based on knowledge or calculation, their performance decreased. Interestingly, Knowledge (β = 0.155, p = 0.268) did not have a significant effect on performance, ruling out the possibility that participants' knowledge rather than intrinsic motivation could be the main predictor of variation in performance.

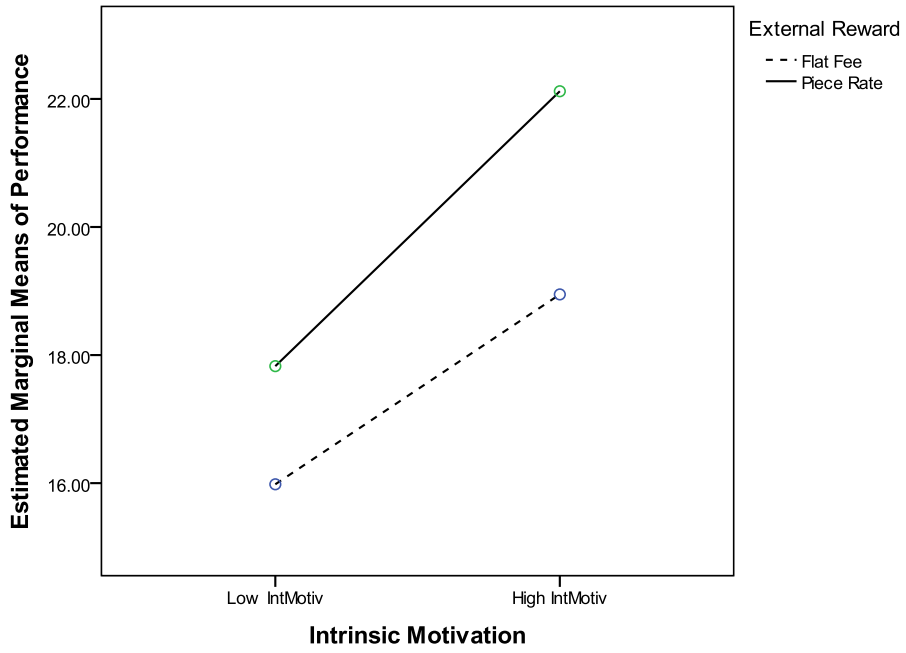
In model 2, we added the three independent variables of IntMotiv, ExtRew, and Field to the first model. As Table 3.4 depicts, both intrinsic motivation (β = 3.688, p = 0.01) and external reward (β = 2.530, $p < 0.05$) significantly improved performance, supporting the second and third hypotheses regarding the positive effect of intrinsic motivation and performance-contingent monetary rewards on performance. The mean performance in the flat fee treatment was 17.86 (N = 44) and the mean performance in the piece rate treatment was 19.48 (N = 48). The effect size of the payment type is 0.25 that matches the results obtained in the past literature on the effect of monetary rewards on motivation and performance (Deci et al., 1999). It indicates that the effect of monetary reward on performance is of small size eventhough it is statstically significant. In addition, the mean performance in the LowIntMotiv treatment was 16.02 (N = 44) and the mean performance in the HighIntMotiv treatment was 21.17 (N = 48). The effect size of intrinsic motivation is 0.87, which indicates that intrinsic motivation has a large effect on performance.

In model 3, we added the interaction between intrinsic motivation and external reward to our model. The multiple regression analysis did not find a significant effect of the intrinsic motivation-external reward interaction on performance (β = 1.328; p = 0.556). However, the data was in the direction that we expected. Figure 3.3 illustrates the interaction effect of external reward and intrinsic motivation on performance. The insignificance of the interaction term can be attributed to our small sample of observations (92 participants). In addition to our main variables, confidence was also found to be significant in the analysis ($ß$ = 1.003, p < .05), indicating that participants who had high confidence significantly performed better than those who had low confidence.

Table 3.4 Multiple regression_depenedent variable: performance

	Dependent Variable: Prformance		
	Model 1	**Model 2**	**Model 3**
(Constant)	16.248 (0.000)	15.691 (0.001)	15.675 (0.001)
Confidence	1.598*** 0.350 (0.001)	0.986* 0.216 (0.043)	1.003* 0.220 (0.041)
PureGuess	-0.089*** -0.371 (0.000)	-0.087*** -0.365 (0.000)	-0.087*** -0.363 (0.000)
Knowledge	0.155 0.107 (0.268)	-0.012 -0.008 (0.931)	-0.014 -0.010 (0.916)
Age	-0.163 -0.123 (0.242)	-0.104 -0.078 (0.432)	-0.110 -0.083 (0.408)
Gender	1.111 0.084 (0.366)	0.422 0.032 (0.735)	0.411 0.031 (0.743)
UnderGrad	0.329 0.025 (0.805)	-0.268 -0.021 (0.829)	-0.153 -0.012 (0.904)
EnglishFirst	-1.564 -0.122 (0.230)	-0.166 -0.013 (0.907)	-0.167 -0.013 (0.906)
Field		3.877* 0.283 (0.013)	3.885* 0.284 (0.013)
IntMotiv		3.688* 0.287 (0.010)	3.629* 0.283 (0.012)
ExtRew		2.530* 0.197 (0.026)	2.509* 0.195 (0.028)
IntMotivxExtRew			1.328 0.052 (0.556)
R Squared	31.6	39.7	44.3
Adjusted R Squared	25.7	32.9	37.2
F statistic	5.341***	6.209***	5.629***
P value	0.000	0.000	0.000
N	88	88	88

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.05, ** significant at 0.01, *** significant at 0.001



Covariates appearing in the model are evaluated at the following values: Confidence = 5.1236, PureGuess = 24.4944, Knowledge = 2.7303, Age = 23.9663, Gender = .6180, UnderGrad = .5618, EnglishFirst = .4719, Field = .6742

Figure 3.3 External reward x intrinsic motivation interaction and performance

In order to test hypotheses 5 and 6 regarding the effect of external reward and intrinsic motivation on overall motivation, we ran multiple regressions with overall motivation as the dependent variable. Overall motivation was assessed by calculating the amount of time spent on the test. Table 3.5 illustrates the results of the step-by-step multiple regression analysis with the behavioral measure of overall motivation as the dependent variable. As for the performance-related models, we entered the covariates into the model in the first step. In the second step, we entered the independent variables of Field, IntMotiv, and ExtRew to our model. Interestingly, the results of the multiple regression analysis indicated that both external reward ($ß = 4.528$, $p < .05$) and intrinsic motivation ($ß = 4.909$, $p < 0.05$) improved overall motivation. Therefore,

hypothesis 5 on the effect of intrinsic motivation on overall motivation is supported. This result supports the implied assumption made by cognitive evaluation theorists regarding the positive effect of intrinsic motivation on overall motivation. Concerning the two competing hypotheses on the effect of external reward on overall motivation, hypothesis 6a was supported. This result is consistent with the behavioral theories of motivation and is in contrast with the cognitive evaluation theory that asserts that performance-contingent monetary rewards undermine motivation.

Table 3.5 Multiple regression_depenedent variable: overall motivation

	Dependent Variable: Overall Motivation		
	Model 1	**Model 2**	**Model 3**
(Constant)	38.004 (0.000)	35.772 (0.000)	35.774 (0.000)
Confidence	-0.442 -0.065 (0.558)	-1.273 -0.188 (0.128)	-1.275 -0.188 (0.130)
PureGuess	-0.033 -0.093 (0.388)	-0.029 -0.083 (0.415)	-0.029 -0.083 (0.417)
Knowledge	-0.184 -0.085 (0.440)	-0.426 -0.198 (0.75)	-0.425 -0.198 (0.077)
Age	0.162 0.082 (0.496)	0.262 0.133 (0.251)	0.263 0.134 (0.255)
Gender	1.164 0.059 (0.578)	-0.263 -0.013 (0.903)	-0.261 -0.013 (0.904)
UnderGrad	-2.906 -0.151 (0.204)	-3.820 -0.199 (0.77)	-3.833 -0.200 (0.082)
EnglishFirst	-2.459 -0.129 (0.269)	0.080 0.004 (0.974)	0.080 0.004 (0.974)
Field		6.806* 0.335 (0.012)	6.805* 0.335 (0.012)
IntMotiv		4.909* 0.257 (0.045)	4.916* 0.258 (0.047)
ExtRew		4.528* 0.237 (0.021)	4.530* 0.237 (0.022)
IntMotivxExtRew			-0.156 -0.004 (0.968)
R Squared (%)	9.7	24.9	24.9
Adjusted R Squared	1.9	15.3	14.2
F statistic	1.237	2.591**	2.326*
P value	0.292	0.009	0.016
N	88	88	88

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.05, ** significant at 0.01, *** significant at 0.001

3.10 Conclusion

In this study we examined the effect of performance-contingent monetary rewards and intrinsic motivation on overall motivation and performance. We developed a new measure for intrinsic motivation by matching the participants' fields of study with the topic of their test. Our new measure resolved the shortcomings of the previous measures of intrinsic motivation, including the confounding of intrinsic motivation with monetary rewards in studies using self-reports and free-choice behavior for assessing intrinsic motivation. The manipulation check supported our hypotheses that self-selection into a specific area of knowledge can be used as a new measure for intrinsic motivation.

The results of this experiment demonstrated important findings: first, performance-contingent monetary rewards positively influenced performance, disregarding the person's level of intrinsic motivation. This finding contrasts with cognitive evaluation theory's prediction that monetary rewards can influence performance only for activities that are not intrinsically motivating (Deci et al., 1999); regardless of the level of intrinsic motivation, participants who received performance-contingent monetary rewards performed significantly better than those who did not receive any rewards. This result supports the behaviorist theories of motivation, indicating that monetary rewards can be used as a motivator for better performance.

Second, the results of the multiple regression analysis on the effect of external reward and intrinsic motivation on the behavioral measure of overall motivation—time spent on the task—show that both external reward and intrinsic motivation can improve overall motivation. The behavioral measure of motivation is considered a more valid and accurate measure of intrinsic motivation than self-reports of interest (Deci et al., 1999); unlike self-reports of interest,

it is unobtrusive; participants do not believe that the experimenter has collected data regarding the amount of time they spent on the task or their free-choice task persistence, so interpersonal factors will have minimal impact on behavioral measures of motivation (Deci et al., 1999).

This result is in contrast with cognitive evaluation theory (Deci et al., 1999) and its supporting experiments (e.g., Pittman et al., 1977), which indicate that performance-contingent rewards damage the behavioral measure of motivation. This difference can be due to the type of behavioral measure that we used in our study and the measures used in these studies for operationalizing motivation.

In this study, we used the time spent on the task as the behavioral measure of motivation while previous studies used time spent in the free-choice period or performance during the free-choice period as a behavioral measure for intrinsic motivation. As was mentioned above in the section on past studies' limitations, studies using a free-choice measure did not consider the amount of time spent on the main task in their analysis, so it is possible that the decrease in the time spent on the task in the free-time period is the result of the participant's increased time spent on the task in the main experimental session; the presence of performance-contingent reward increased motivation during the experimental sessions, which could have resulted in fatigue or a satiation effect and a subsequent decrement in time spent on the task in the free-time period (Calder & Staw, 1975a). Our results partly support this argument by demonstrating that the introduction of performance-contingent rewards during the experimental session increased participants' motivation.

Furthermore, our findings have important implications for research in motivation. They indicate that performance-contingent rewards can improve both motivation and performance.

Performance improvement is one of the ultimate goals of any organizational setting; our findings support the widespread use of performance-contingent rewards as a motivational strategy in applied settings.

However, our study has some limitations. First, we conducted our experiment using a group of students at the University of Calgary. Further examinations in other university and non-university settings can help in the generalization of these findings.

Second, we only tested the effect of a moderate level of performance-contingent pay—$0.50 per correct answer—on motivation and performance. Testing other extreme contingent payment levels such as very high and very low pay can give better insight into the interaction effect of performance-contingent payment and intrinsic motivation on performance.

In later studies, we may change the level of performance-contingent payment from very low to very high—for example, we may use variable payments ranging from $0.10 to $1—and see how it may affect overall motivation and performance. While a moderate level of contingent pay can improve performance as supported in this study, there is a possibility that a very small or very high level of performance-contingent pay will decrease performance. When the compensation amount is less than the amount of money expected by the participant for the time and effort allocated to the task, subjects feel insulted and decrease their performance as compared to no-pay condition (Gneezy & Rustichini, 2000). On the other hand, when the compensation level is very much higher than expected by the participant for the time and effort allocated to the task, performance may diminish due to the psychological mechanism known as the “choking under pressure” phenomenon (Ariely et al., 2009; Baumeister, 1984).

Finally, our study only focused on the interaction between a special type of payment scheme, performance-contingent pay, and intrinsic motivation on motivation and performance, so the findings cannot be generalized to other types of payment such as completion-contingent and task-noncontingent ones. Using our new measure for intrinsic motivation in future studies, we can examine the interaction effect of other payment schemes and intrinsic motivation on motivation and performance.

Chapter Four: Literature Review on Referral Behavior in Health Care Systems

4.1 Introduction

Variation in referral rates and referral appropriateness among general practitioners (GPs) has always been a prime focus of interest in healthcare research. The interest is due to three related concerns: first, studies on referral rates have confirmed the existence of a systematic variation in rates of referral between general practices and between individual physicians (McBride et al., 2010; O'Donnell, 2000; Hutchison, 1993; Wilkin, 1992; Coulter, 1998; Wilkin & Smith, 1987a; 1987b); second, referral to specialists is related to the use of expensive and finite health care resources. Thus, referral rates variation has financial implications for health care systems. In a study by Crombie and Fleming in 1988, for a practice of approximately 2000 patients the hospital expenditure associated with the lowest and the highest rates of referral varied between 40000 and 408000 Euros—which showed a 10-fold difference in expenditure based on the referral rate; finally, variation in referral rates creates concern with regard to equitable patient care. If some doctors refer too many patients while others refer too few patients, some patients may receive unnecessary procedures and investigations while others may be denied necessary specialist care (Newton et al., 1991).

Previous studies have shown that variation in referral rates can be explained by both GP-unrelated factors (Giuffrida et al., 1999; Reid et al, 1999) and GP-related factors (O'Donnell, 2000). Based on the referral literature, factors explaining the variation in referral rates fall into four categories of: patient characteristics, practice characteristics, GP characteristics, and payment systems (O'Donnell, 2000). In the following sections, we will review the results of the

studies on the effect of each of the mentioned categories on referral patterns of general practitioners.

Factors including GP-specialist relationship and payment schemes will be the focus of our next two experiments. In chapters four and five, we will explain the results of the two experiments that we conducted in order to examine the effect of GP-specialist relationship and payment schemes on referral patterns respectively. For the purpose of these studies, physicians were randomly assigned to different treatment conditions based on the type of the relationship between the general practitioners and the specialists and the type of physicians' payment schemes. In each treatment, they were presented with a set of hypothetical cases in two areas of rheumatology and respirology. The cases were developed based on real patients who had been referred to specialists in these areas. We selected a subset of the cases based on a panel of specialists' opinion regarding the rate of referral urgency and whether or not the case should be referred to a specialist.

4.2 Patient Characteristics

Patient demographics such as age, gender and socio-economic status (SES) may explain part of the variation in referral rates among GPs. In two studies adjusting the referral rates for the patients' age and sex reduced the observed variation by less than 10% (Roland et al., 1990a; Morrell et al., 1971). A quantitative study in UK showed that referral rates increased as the patient's economic status improved. Other patient characteristics that influenced referral rate were patient gender and age: for example, GPs were less likely to refer women than men for hip pain (McBride et al., 2010). In addition, GPs referred older patients—specifically women with

postmenopausal bleeding and patients with hip pain—less than younger patients (McBride et al., 2010).

As other important patient characteristics influencing referral rate, patient concerns, reassurance, expectations, needs and values, and pressure can be mentioned (Newton et al., 1991). In one study conducted in US, 13.6% of referrals were due to patient pressure (Forrest et al., 1992). Another qualitative analysis of patients with headache showed that patient anxiety and pressure led the GPs to refer patients to specialists due to frequent primary care visits, ineffective communication between the GP and the patient, and time constraints (Morgan et al., 2007).

Since we used the same set of hypothetical cases in all treatments in each of the two experiments, patient characteristics and its effect on referral was not addressed in these studies.

4.3 Practice Characteristics

General practitioners work in various work settings including solo offices, group offices, unique primary care teams, walk-in clinics, academic health centers, and emergency departments. There is conflicting evidence on the relationship between practice size and variation in referral rates (O'Donnel, 2000). Some studies found no relationship between practice size and referral: one study comparing high and low referring GPs found no significant difference in their list size or number of partners (Wilkin & Smith, 1987b); another study—in Lincolnshire—found no difference in referral rates between single-handed GPs and GPs in partnership (Madeley et al., 1990).

However, other studies found either a positive or negative relationship between practice or list size and referral rate: a study in Nottinghamshire found a positive relationship between single-handed practices and referral rates (Hippisley-Cox et al., 1997); another study in Denmark

found a significant negative relationship between practice size and referral rate; referral rates fell as practice size increased (Christensen et al., 1989); finally, a study in the Netherlands found that referral rates increased as the GPs' list size (Delnoij et al., 1997; Kerssens & Groenewegen, 1990) or the number of GPs in the practice increased (Verhaak, 1993).

In addition to practice size, geographical location may also influence referral rates; higher referral rates were associated with shorter distances from the practice to the outpatient clinic (Jones, 1987). Several studies both in England and Canada showed that rural GPs had significantly lower referral rates in comparison with urban GPs (Langley et al., 1997; Madeley et al., 1990): the difference was mostly due to nonmedical factors such as accessibility to specialists, lab tests, beds and other resources (Langley et al., 1997). In order to test the effect of geographic location on referral rate in our experiments, we requested physicians to indicate the location of their practice and examined its effect on referral rate.

An American study showed that 75% of American rural counties were medically underserved due to the lower socio-economic status of the patients in these areas (Ryan-Nicholls & Racher, 2004). Similar to the United States, a province-wide study in Canada found variation in referral rates based on geographical location: the variation could be attributed to patients' ethnicity and socioeconomic status (Sibley & Weiner, 2011). This study also showed differences in referral rates between Canadian provinces—for example, Quebec patients had more visits to specialists and fewer visits to GPs compared to other provinces. The difference was due to the fact that self-referral was acceptable in Quebec unlike other provinces where referrals could only be done through primary care physicians (Sibley & Weiner, 2011).

4.4 GP Characteristics

GP characteristics can influence different aspects of medical decision making (Mckinlay et al., 2002; Wilkin & Smith, 1987a; 1987b). In one study, McKinley et al. found a significant relationship between GP characteristics (e.g. area of specialty, age, and race) and their clinical decisions—such as the most likely diagnosis, level of uncertainty adhering to that diagnosis, and the number of lab tests that would be ordered (McKinley et al., 2002).

As regards with referral decision making, the effect of GP characteristics on referral decisions has not been very clear and empirical evidence is inconsistent. Many researchers have argued that a GP's age can be a possible source of referral rate variation (Wilkin & Smith, 1987a). For example, Evans and McBride stated that older physicians may use lower levels of hospital services than younger physicians (Evans & McBride, 1968). But there is little empirical evidence to support this hypothesis. While two studies in Finland (Vehvilainen et al., 1996) and the U.S. (Bachman & Freeborn, 1999) supported this hypothesis by finding that young, inexperienced GPs had significantly higher referral rates in comparison with older GPs, other studies did not find support for this hypothesis: the Wilkin and Smith study, for example, showed that doctors with high referral rates comprised a greater proportion of more experienced doctors compared to doctors with low referral rates (Wilkin & Smith, 1987b); although the difference was not statistically significant, this finding was in contradiction with the hypothesis that older, more experienced doctors refer less than younger, inexperienced doctors. In another study, Franks et al. found that physicians who had more years in practice had higher referral rates (Franks et al., 2000).

In fact, most studies on the effect of age on referral rate found no significant results. As an example, one study conducted by Wright on 68 doctors showed no relationship between age and referral rate (Wright, 1968); another study on 369 general practitioners showed similar insignificant results on the relationship between age and referral (Forsyth & Logan, 1968). Similar to these two studies, another study in UK found no relationship between GPs' referral rates and their age and years of experience (Cummins et al., 1981).

Physician gender is another factor that might affect referral rate. Some studies done both in the United States (Forrest et al., 2006; Bachman & Freeborn, 1999) and England (Wilkin & Smith, 1987b) found no difference between male and female physicians in their referral rates. However, a study done by Franks et al. (2000) found that female physicians were more likely to refer patients (Franks et al., 2000).

GP knowledge and interest in a particular area of specialty is considered another predictor of referral rate (Newton et al., 1991; Morrell et al., 1971; Evans & McBride, 1968). Having particular interests in a specific area of medical specialty can have a mixed effect on referral rates. On the one hand, interest in a specific area may decrease the referral rate since it may result in the physician's higher level of knowledge, experience and confidence in that area (Newton et al., 1991). On the other hand, it may increase referral rate in that specific area since it may attract more patients to come and visit the GP for that particular specialty (Newton et al., 1991; Morrell et al., 1971; Evans & McBride, 1968). One study on five GPs showed that GPs with specialties in ear, nose, and throat and ophthalmology had higher referral rates to specialists in these areas. This effect persisted after adjusting for case mix. The high rate of referral could not be explained by GPs' lack of confidence since these GPs felt more confident than average in

managing cases in this specific area (Reynolds et al., 1991). Another study by Evans and McBride showed that doctors with a special interest in pediatrics and rheumatology referred more patients to specialists in these areas (Evans & McBride, 1968).

A study done in Alberta, Canada showed a difference in referral patterns between physicians educated in Canada and physician who studied internationally and now practice in Canada. International medical graduates (IMGs) found the referral process a long and difficult process which affected their workload and the scope of their practice. They also insisted on the need for acquiring tacit knowledge about which specialists to consult and which patients to send to emergency (Lockyer et al., 2007).

The GP-specialist relationship is another determinant of referral rate and referral appropriateness in decentralized referral systems (Hajjaj et al., 2010; Langley et al., 1997; Newton et al., 1991). In traditional decentralized referral systems, GPs refer the patients directly to the specialists whom they know. Knowing the specialist can influence GP's referral decision making about cases which fall into the uncertain category in different ways: first, it can facilitate an informal contact between the GP and the specialist. A doctor can call a specialist to obtain an opinion about the appropriateness of referring a problematic case (Newton et al., 1991); second, such social relationships can motivate the GPs to be more careful about their referrals. The GP is mindful of not wasting specialists' time and may feel guilty about referring patients who they might be able to manage themselves (Newton et al., 1991). Thus, social relationships between GPs and specialists can result in more effective GP-specialist communication which in turn can lead to higher quality referral processes (Gandhi et al., 2000). However, it is also possible that the presence of social relationships between the GP and specialist would increase the referral

rate, since GPs may feel more comfortable to send patients to a specialist whom they know. In one study, Langley et al. did not find a conclusive result on the impact of GP-specialist relationship on referral rate. There was a significant variation among different groups of GPs about the effect of GP-specialist relationship on referral rates (Langley et al., 1997). The basis for the definition of the GP groups was their geographic location. GPs of different geographic locations showed significantly different perspectives on the effect of GP-specialist relationship on referral rate. In order to experimentally examine the effect of GP-specialist relationship on referral rate, we designed and conducted an experiment to investigate the effect of GP-specialist relationship on referral rate in a centralized referral system versus a decentralized referral system, characterized by a close relationship between the general practitioners and the specialists. Chapter 4 explains the design and results of this experiment.

Referral variation among general practitioners often persists even after considering physicians' background, experience, and practice and patient characteristics (Wilkin & Smith, 1987a). A high level of unexplained variance remains even after controlling for all these factors. The existence of unexplained variance suggests that referral decision making may be subject to psychological factors that cannot be predicted by strictly rational or utilitarian decision making principles; a GP's tolerance for uncertainty (Morrell et al., 1971), risk seeking preferences, (Newton et al., 1991; Holtgrave et al., 1990; Nightingale, 1988; 1987a; 1987b) and fear of litigation (De Marco et al., 1993) are examples of such psychological factors that may influence referral decisions (Slovic & Lichtenstein, 1983).

Morrell et al. suggested that doctors' referrals could be attributed to their perception of the need for specialty care, which is in part related to their tolerance for uncertainty (Morrell et

al., 1971). Cummins et al. stated that doctors may have unique referral thresholds that are characterized by all their personal characteristics such as training, experience, uncertainty tolerance, sense of autonomy and personal enthusiasm (Cummins et al., 1981).

Making a referral to a specialist, ordering a lab test, and referring a patient to hospital are different strategies that physicians commonly choose to avoid or decrease risk and uncertainty (Nightingale, 1987b). Therefore, physicians' risk preferences can highly influence the type of strategies physicians choose for their patients. Physicians who are risk averse may order more lab tests (Nightingale, 1987a; 1987b) or refer more patients to specialists (Holtgrave et al., 1990) in order to avoid uncertainty or any possible risks such as the risk of the patient's worsening health conditions or the risk of probable litigation. However, physicians who are risk seeking may refer significantly lower number of patients to specialists due to their higher tolerance for uncertainty and less fear of related risks.

In a series of interesting empirical studies, Nightingale found that physicians who were risk seeking in losses—loss averse physicians—ordered significantly more laboratory tests both in their general practice clinics (Nightingale, 1987a) and for hypothetical patient cases (Nightingale, 1987b), referred significantly more patients to hospital emergency rooms (Nightingale, 1988), had a higher preference for intubation (Nightingale & Grant, 1988), and resuscitated patients significantly longer (Nightingale & Grant, 1988). Holtgrave et al.'s study replicated these results on the effect of risk attitude on laboratory use. GP's risk attitude explained over 50% of the variance in use rate of several laboratory procedures (Holtgrave et al., 1990).

However, risk attitude (Franks et al., 2000; Holtgrave et al., 1990), tolerance for uncertainty, and fear of malpractice (Franks et al., 2000) did not have a significant association with referral. Tolerance for uncertainty had a positive but weak association with referral rate (Franks et al., 2000, Bachman et al., 1999). Greater reluctance to disclose uncertainty to patients and less reluctance to disclose uncertainty to other physicians were found to be associated with referral (Forrest et al., 2006).

4.5 Payment Systems

Economic theories and common sense both suggest that payment methods can influence performance (Conrad & Christianson, 2004; Bull et al., 1991; Donaldson & Gerard, 1984; Evans, 1974). Research outside health care has found that financial incentives can have a positive effect on employee motivation and performance. But, a meta-analysis showed that this positive effect is not always guaranteed—specifically in complex systems that need careful design and integration within the organization (Glasziou et al., 2012; Jenkins et al., 1998).

In healthcare, the existence of the strong system of ethics may decrease or even completely eliminate the effect of payment systems on doctors and prevent them from providing ineffective services in order to increase their own personal income (Gosden et al., 2001). However, empirical evidence shows that different types of payment systems have influenced physician performance (Devlin & Sarma, 2008; Gosden et al., 2001; Gosden et al., 1999; Hughes & Yule, 1992; Goldmann, 1952).

An overview of four systematic reviews in healthcare shows that the effect of payment on performance is inconsistent and unclear (Flodgren et al., 2011). Financial incentives had a mixed effect on consultation or visit rates. They generally improved processes of care (such as

diagnostic and curative services, prescriptions, healthcare utilisation, etc) and referrals and admissions. In addition, they successfully reduced prescribing costs. However, financial incentives proved ineffective in improving guideline compliance (Flodgren et al., 2011). Another review of seven studies in primary care found that financial incentives were useful in improving specific outcomes in specific settings. However, the study concluded that there was insufficient evidence to advocate or abandon the use of financial incentives to improve the quality of primary health care (Scott et al., 2011).

The main methods of remunerating physicians are capitation, salary, fee-for-service, fundholding, and pay-for-performance. In the following sections, we will define each payment method's characteristics and examine its effects on GP referral behavior. Our literature review indicated that previous studies had not tested the effect of different payment methods on physicians' referral behavior in a controlled experimental setting. In order to address this issue, we conducted an experiment and examined the effect of two types of payment schemes, including fundholding and pay-for-performance on referral rate and referral appropriateness in general practice. Chapter 5 explains this experiment.

4.5.1 Fee-for-Service

Fee-for-service (FFS) is one of the most common types of GP remuneration in many countries including Canada (Holden & Madore, 2002). In a fee-for-service payment scheme, physicians get paid for each service they provide to patients according to a pre-determined schedule of tariffs. Therefore, physicians' annual income depends on the quantity and type of health services provided and the level of pre-determined fees (Holden & Madore, 2002; Skedgel, 1996).

From the perspective of cost coverage, fee-for-service is a simple and transparent method of remuneration. It can be applied to all physician practices disregarding its size or type of practice. As a result of this simplicity, a lot of countries—including Australia, Canada, Japan, Germany, and Belgium—use fee-for-service as one of their main methods of GP remuneration (Holden & Madore, 2002). In Canada, most primary care physicians work in solo or group practices. So, they are responsible for covering the costs associated with delivering health care services including costs of administration, supplies and staff. Since these costs are all considered in the pre-set fee schedules, fee-for-service is a suitable method for the Canadian health care system (Holden & Madore, 2002).

Compared to other systems, the fee-for-service payment system has several advantages: from the patients' perspective, patients have freedom of choice in selecting their doctors, and they are also able to change their doctors or seek a second opinion if they wish (Holden & Madore, 2002); from the physician's perspective, FFS is beneficial since physicians do not bear the risk of service-intensive patients in this system. Fee-for-service payment schemes shift the financial risk associated with less-healthy patients away from physicians to insurers (Devlin et al., 2006); from the perspective of quality of service, studies found that FFS could produce higher quality of care (Tarlov et al., 1989) and better access to health services (Helfinger & Northrup, 2000; Brudevold et al., 2000) as compared to capitation payment. In addition, FFS physicians conduct more patient visits (Devlin & Sarma, 2008) and allocate more hours of direct patient care in the office and clinic (Sarma et al., 2010) compared to non-FFS physicians. Fee-for-service can also foster productivity among GPs since the physicians get paid according to their workload (Eisenhardt, 1989; Porter & Lawler, 1968). The FFS payment method has

produced higher satisfaction among both patients and physicians (Brudevold et al., 2000; Nadler et al., 1999) due to its direct relationship between the quantity of work and outcome (Conrad & Christianson, 2004).

However, fee-for-service payment method may have disadvantages also: it may encourage physicians to over-provide services to patients (Tu et al., 2009; Evans, 1974). When the marginal cost of providing a unit of service is less than the marginal revenue, doctors can increase their revenue by increasing the number of services they provide. Physicians may provide more services when there is uncertainty about appropriate treatment or when they face reduced income due to the small number of patients or changes in the number of physicians or their fee levels (Cromwell & Mitchell, 1986; Woodward & Warren-Bolton, 1984; Yett et al., 1983). In practice, FFS payments are generally above the marginal costs so, the payment covers the administrative and fixed costs associated with physician's practice, creating financial incentives for physicians to provide more services to patients than they would do in an incentive-neutral system (McGuire, 2000; Evans, 1974).

Empirical studies provide evidence on the over-provision of services under FFS payment systems. A literature review by Gosden et al. (2001) showed that FFS increased the use of primary care services in comparison with other forms of physician payment. Healthcare utilization, defined as per-capita consultations, was found to be higher under the FFS system rather than capitation or blended payment systems. Healthcare utilization was twice as much in FFS systems as in salary-based systems. In another randomized controlled study, FFS led to a 22% increase in utilization rate in comparison with salary-based system (Hickson et al., 1987). In

a study in Newfoundland, FFS payment showed a strong association with higher antibiotic prescription rates as compared to salary payment (Hutchison & Foley, 1999).

The effect of FFS on referral rate has been mixed based on the empirical literature. It is generally argued that fee-for-service payment may result in low rates of referral by inducing over-treatment among general practitioners (Gosden et al., 2001). In a study in Denmark, researchers changed the payment scheme for a group of physicians from a capitation method to a mixed FFS/capitation system. The study showed that referrals to specialists and hospitals (which were not paid for by fees) decreased significantly 12 months after FFS was introduced to the capitation group compared to the control group (Krasnik et al., 1990). In another study, Davidson et al. evaluated the effects of a change in remuneration system from low cost FFS to either high cost FFS or capitation-based payment with some degree of risk sharing by the provider of secondary care for the management of Medicaid eligible pediatric care (Davidson et al., 1992). Contrary to the results of Krasnik et al.'s study, they observed that the change from FFS to a capitation-based payment resulted in a decrease in referral rates for non-primary care services. However, the change from low cost FFS to high cost FFS had little effect on referral rate (Akbari et al., 2008; Davidson et al., 1992).

In addition to higher utilization and lower referral rates, FFS may result in the provision of services to patients regardless of their effectiveness or necessity (Rosen, 1989); In certain circumstances, FFS may lead to excessive and unnecessary provision of health services, a phenomenon called supplier-induced demand: physicians provide more services to patients than patients would demand if they had the same level of information. While there is a large body of literature on supplier-induced demand in health care (Carlsen & Grytten, 2000; De Jaegher &

Jegers, 2000; McGuire, 2000; Sorensen & Grytten, 2000; Evans, 1971), the empirical results are mixed. Several empirical studies have reported a positive association between FFS payment and supplier-induced demand among physicians (Hickson et al., 1987; Tussing & Wojtowycz, 1986, Rice, 1983). However, many studies failed to support the supplier-induced demand hypothesis (Grytten & Sorensen, 2001; Sorensen & Grytten, 1999). Some researchers have argued that supplier-induced demand and difference in utilization rate under FFS systems may be influenced by supply factors such as the number of available hospital beds and physician density (Carlsen & Grytten, 2000; De Jaegher & Jegers, 2000). The seemingly contradictory empirical evidence found in the empirical studies could be attributed to excess supply of physicians in relation to need in different geographical locations (Devlin et al., 2006).

In general, FFS payment schemes are expensive to implement due to the need for a large infrastructure and the number of workers to keep item-wise records on the volume and type of services for payment purposes. High administrative costs, over-utilization of services, and the existence of supplier-induced demand in FFS systems may make FFS sysem an expensive option in comparison with other payment schemes (Devlin et al., 2006).

4.5.2 Capitation

In capitation systems, the GP receives a flat fee-per-patient payment for each registered patient and is obliged to provide specific services to these patients. GPs are paid a predetermined amount for each patient registered or enrolled in their care. The amount of the pre-set capitation fee for each patient is calculated based on the current patterns of average annual use of primary care services across the entire population and is adjusted according to patients' characteristics

known as risk-adjusting factors such as age and sex. The fee for high-user patients is greater than the fee for low-user patients. This fee may cover all or some services a patient may receive.

Physician capitation payments should be distinguished from capitation payment made to health care organizations such as health maintenance organizations in the USA and GP fundholders in the UK. Under a capitation system, physicians are paid a fixed fee for each patient regardless of the patient's health status or the number of patient visits.

In capitation payment systems, patient visits are a source of expense to physicians as opposed to a source of revenue. Thus, although capitation eliminates the problem of over-providing health services, it creates an incentive to under-provide health services to patients. Unlike FFS physicians, capitation physicians bear the financial risk of serving less healthy patients. Therefore, capitation payment may encourage physicians to select low-risk patients or actively discourage high risk patients unless the fee is adjusted upwards for high risk patients (Gosden et al., 2001). This act of selecting special types of patients is known as cream-skimming (Hausman & LeGrand, 1999; Matsaganis & Glennerster, 1994; Ellis & McGuire, 1986)

Physician behavior in a capitation payment system is highly dependent on the amount of capitation payment. If the fee is smaller, so that it covers only primary care services, there is an incentive to refer the patient to the specialist. In this case, capitation can increase rates of referral to both specialists and hospitals (Hughes & Yule, 1992; Goldmann, 1952). However, if the fee covers both primary care and hospital services, then physicians will have an incentive to choose the cheapest option (Lerner & Claxton, 1994). Capitation may also encourage physicians to withhold care, resulting in under-treatment (Eliss & McGuire, 1986). However, the incentives to

contain costs will be counteracted since physicians in capitation systems need to attract and maintain patients to maximize their income (Hausman & LeGrand, 1999).

Although capitation payment may have negative effects on physicians' behavior, it may have positive effects also. As an example, capitation may encourage GPs to promote long-term preventative health care solutions such as physical fitness or a healthy diet in order to reduce the likelihood of future patient visits (Goldmann, 1952).

4.5.3 Salary

Salary is the simplest type of payment. In salary-based systems, GPs receive a flat annual income regardless of the number of patients they visit or the volume of services they perform (Holden & Madore, 2002). Salary payment provides an incentive for physicians to minimize their personal costs. As a result, physicians may select low risk patients, write prescriptions, or refer patients to specialists (Grytten et al., 1995). Thus, from the perspective of referral rates, salary-based payment may result in high referral rates to specialists since it provides no financial incentives for GPs to increase the level of services provided for their patients.

Salary-based payment has several advantages. Since under salary payment, physicians receive a fixed annual payment, they will have no incentive to over-supply services, prescribe unnecessary medication, or encourage unnecessary visits (Holden & Madore, 2002). In salary-based payment, doctors have more time to devote to their patient consultation or to engage in preventative and long-term care. To some extent, salary-based payment may result in more efficient use of health care services. Furthermore, the salary type of payment contains no financial risk for the physicians. GPs are paid regardless of the number of patients or the type of

services they provide. In general, similar to capitation type of payment, salary payment may result in under-treatment, while fee-for-service may result in over-treatment.

4.5.4 Fundholding

Fundholding is considered to be one of the major types of payment reforms in health care systems. Under the fundholding policy, GPs take on budgets for purchasing elective (planned) care from hospitals and/or other health care providers (Brereton & Vasoodaven, 2010). In the 1990s, fundholding policy was deemed to have been one of the most promising reforms in terms of improving secondary care provider quality and responsiveness (Brereton & Vasoodaven, 2010). In the English fundholding scheme, which was effective from 1991/1992 to 1998/1999, GPs could elect to receive a budget to meet the costs of certain types of elective surgery (chargeable electives) for their patients.

Fundholders were usually given relatively generous budgets. They did not pay for non-chargeable electives or for emergency admissions. Their budget deficits were often covered by Health Authorities (HAs), and there were limits on the maximum cost per period borne by fundholders (Dixon et al., 1994; Audit Commission, 1996). They were able to retain any budget surplus. However, the surplus was not intended to be part of GP income. It was required to be spent for the benefit of the practice's patients by providing additional services, new equipment, etc. Practices could use the surplus to buy additional services for themselves. Since many GPs own their practice accommodation, they could also benefit from investments that increased its value. Non-fundholding practices did not bear the cost of any type of hospital admissions. Fundholding was abolished in April 1999. However, a new version of the policy was reintroduced in England in April 2005 (Dusheiko et al., 2006). In this version, practices were

given a budget to cover all admissions, not just electives, and were permitted to keep any savings and use it for the benefit of their patients (Department of Health, 2004).

The main objective of fundholding policy was to encourage GPs to manage their budgets more appropriately. It was expected that fundholding policy would provide the necessary incentives for GPs to manage their prescribing and referral patterns more effectively and efficiently and to commission care that better reflect the requirements of their local populations (Dixon & Gelennerster, 1995). However, empirical studies regarding the effect of fundholding on different aspects of health services shows mixed results. In some studies, researchers found that fundholding policy resulted in a 0-24% reduction in prescription costs (Dowell et a., 1995) and in the total number of drugs per prescription (Himmel et al., 1997; Wilson et al., 1996; Whynes et al., 1995; Maxwell et al., 1993; Bradlow & Coulter, 1993). In another study, Redfern & Bowling (2000) found little difference in waiting time and patient satisfaction between fundholding and non-fundholding practices (Redfern & Bowling, 2000).

As regards the effect of fundholding on referral patterns, some studies showed that fundholding produced positive results and decreased cost of service. For example, Surender et al. compared the referral rate of fundholder GPs and non-fundholder ones before and after the introduction of the fundholding policy. The results of this study showed that the referral rates of non-fundholders increased by 26.6 percent after three years while the referral rate of fundholders had increased by only 7.5 percent within the studied three-year interval (Surender et al, 1995). Another study showed that a shift from a fee-for-service payment system to fundholding reduced the number of referrals for elective surgery and to private clinics (Chaix-Couturier et al., 2000). Dusheiko et al. (2006) found that referral rates for chargeable elective admissions increased for

ex-fundholders after the abolishment of the scheme (Dusheiko et al., 2006). This result can be due to the fact that GPs may have been conservative with referrals while they were fundholders in order to maximize savings.

Contrary to the results of the above-mentioned studies, Coulter and Bradlow (1993) found no significant difference between the referral rates of fundholders and non-fundholders. They compared referral rates from 10 "first wave" fundholders with 6 non-fundholding practices during the preparation year (phase 1) and one year after the introduction of the fundholding scheme (phase 2). The fundholding budget for the first year was based on the referral rate in the preparatory year. After the initiation of fundholding reforms, both groups of practices increased their referral rates. The fundholding practices increased their referrals from 107.3 per 1000 patients per annum (95% confidence interval 106 to 109) to 111.4 (110 to 113) and the non-fundholders from 95.0 (93 to 97) to 112 (110 to 114). No difference in overall referral rates was found between fundholders and non-fundholders in this study (Coulter & Bradlow, 1993). There was a significant increase in the non-fundholders referral rate from pre-intervention to post-intervention period (median pre fundholding 95.9 annual referrals per 1000 patients versus median post-fundholding 117.2 annual referrals per 1000 patients) (Akbari et al., 2008). Based on Akbari et al., this study has unit of analysis error and provides insufficient data that makes re-analysis of its results impossible.

Surender et al.'s study was in fact a follow-up of Coulter and Bradlow's study (1993) in order to test the effect of fundholding on referral rate for a third time period. However, by the time of this study four non-fundholding practices had become fundholders or shadow fundholders (Akbari et al., 2008). In another study, Kammerling & Kinnear (1996) evaluated the

effect of fundholding on outpatient referrals to orthopedic clinics one year before becoming a fundholder and two years after that. In the year before fundholding, the practices in the fundholding group referred fewer patients than the control group although the difference was not significant. However, two years after the fundholding program, fundholders' referral rate increased by 13% while non-fundholders referral rate increased by 32%. Therefore, fundholders increased their referrals less than non-fundholders two years following the prgram (Kammerling & Kinnear, 1996). However, in their review of the factors influencing referral rate from GPs to specialists, Akbari et al. (2008) states that the Kammerling and Kinnear's study suffers from a unit of analysis error and its authors did not report the statistical significance of their findings (Akbari et al., 2008).

Based on the literature on fundholding, it can be concluded that the results of studies on the effect of fundholding on referral are inconclusive and not generalizable since these studies have tended to be small scale, lack adequate controls, and have methodological limitations (Dusheiko et al., 2006). Therefore, further analysis in a controlled setting is still needed in order to get a better understanding of the effect of fundholding on referrals.

4.5.5 Pay-for-Performance

Pay-for-performance is a type of payment method that directly relates a proportion of the remuneration of physicians and health care providers to the achieved results on quality factors of their performance. Pay-for-performance schemes are increasingly used to enhance the quality of care, including the referral patterns of GPs (An et al., 2008; Srirangalingam et al., 2006; Grady et al., 1997). Despite some researchers' skepticism about the effectiveness of pay-for-performance schemes in improving health care quality (Rosenthal & Frank, 2006), studies on the relationship

between pay-for-performance schemes and the quality of care have found that pay-for-performance contracts do affect physician behavior and improve the quality of provided primary care services (Van Herck et al., 2010; Campbell, 2007; McElduff et al., 2004; Gosden et al., 2001; Chaix-Couturier et al., 2000). Quality goals and targets as addressed in these studies, fall into three categories of process indicators (Herrin et al., 2008; Lindenauer et al., 2007; Grossbart; 2006; Chung et al., 2003), intermediate measures (Weber et al., 2008; Tahrani et al., 2007; Beaulieu & Horrigan, 2005; Larsen et al., 2003;), and outcome measures (Bhattacharyya et al., 2009; Ryan, 2009; Downing et al., 2007; Twardella & Brenner, 2007). While intermediate and outcome measures primarily focus on health outcomes, process indicators refer to clearly defined measures that should be taken to improve the quality of care in health care systems. In one study, Grossbart (2006) used a set of 17 process measures to calculate a composite quality score for process improvement in three clinical areas, including acute myocardial infarction (AMI), heart failure, and pneumonia. As examples of process measures related to AMI, the use of aspirin at arrival and at discharge, smoking cessation advice/counseling can be mentioned (Grossbart, 2006).While early pay-for-performance programs generally focused on improving the quality of care for one specific patient group (e.g. immunization), recent pay-for-performance programs have increased the number of patient groups and quality targets covered by their program (Van Herck et al., 2010).

Pay-for-performance schemes are necessary in health care because the common methods of remuneration do not reward physicians for higher quality as do prices in most other markets (Robinson, 2001). This effect is often complex and limited (Christianson et al., 2008) and is dependent on the pay-for-performance design choices and the context (Van Herck et al., 2010).

For example, Van Herck et al.'s (2010) review found that programs that have focused on process improvement resulted in higher improvement rate than those that focused on intermediate and outcome measures.

Van Herck et al. (2010)'s review is one of the most recent reviews on pay-for-performance in health care. Many reviews were done on the effect of pay-for-performance programs on health care systems prior to Van Herck et al.'s study. These reviews had identified a dearth of studies on pay-for-performance programs (Frolich et al., 2007; Rosenthal & Frank, 2006; Armour et al., 2001; Dudley et al., 2004; Kane et al., 2004). However, as can be witnessed in the Van Herck et al.'s study, the number of studies has increased dramatically over the last 20 years. The study done by Van Herk et al. adds 79 more studies to the previously examined studies and includes two additional years as compared to the previous reviews (Mehrotra et al., 2009; Christianson et al., 2008; Sabatino et al., 2008; Schatz, 2008; Conrad & Perry, 2009; Greene and Nash, 2009).

Van Herck et al.'s review attempted to find how pay-for-performance schemes used in health care settings affected different clinical factors such as effectiveness, access, equity, coordination, patient-centeredness, and cost-effectiveness. With regard to the clinical effectiveness and quality improvement of the pay-for-performance programs, the results of the review showed that the clinical effects ranged from negative or absent to positive (1-10%) or very positive (above 10%), depending on the type of target and the type of the program (Van Herck et al., 2010). The majority of studies showed positive results. Negative results were only found in a few studies. Three studies had negative results on only one target while they had positive results on other targets (Mullen et al., 2009; Pearson et al., 2008; Grossbart, 2006).

However, it should be noted that negative results in this context relates to less quality improvement in pay-for-performance payment systems in comparison with non-pay-for-performance systems. It does not mean a decline in quality under the pay-for-performance programs. In general, the use of pay-for-performance programs resulted in a 5% quality improvement, but the results showed much variation depending on the measure and the program.

According to the reviewed studies, pay-for-performance programs can improve specific targets when design choices and context are optimized and aligned. The effect of the payment program on performance depends on the primary mission of the program. The majority of studies showed that when the program supports uniform minimal standards, it is able to serve its purpose successfully. However, when the program is intended to boost performance of all providers, its capability becomes limited and is confirmed for only a number of targets, such as for diabetic care.

Another interesting finding of this review was related to the impact of different types of incentives on performance. Positive financial rewards (de Brantes & D'Andrea, 2009; Rosenthal et al., 2008; Coleman et al., 2007; Fairbrother et al., 2001; Fairbrother et al., 1999; Hillman et al., 1998; Morrow et al., 1995) proved to generate more positive effects on performance than incentives that were based on a competitive approach in which there were both winners and losers (Mullen et al., 2009; Karve et al., 2008, Lindenauer et al., 2007; Glickman et al., 2007; Young et al., 2007; Levin-Scherz et al., 2006; Morrow et al., 1995). However, the relationship between financial reward and performance was not linear and was influenced by other factors such as incentive size and level of stakeholder involvement.

Regarding cost effectiveness, one study showed a 2.5-fold return on investment in designing the pay-for-performance program as a result of cost saving in this program (Curtin et al., 2006). Another study on the Premier project in the US showed that the pooled resources collected from penalties were smaller and not sufficient to cover bonus expenses in the project (Kahn et al., 2006). Further, four other studies confirmed the cost effectiveness of the programs (Salize et al., 2009; An et al., 2008; Nahra et al., 2006).

Two types of reward options were fixed thresholds and continuous scale, which produced positive effects in some UK studies but mixed effects in other ones (Herrin et al., 2008; Fairbrother et al., 1999; Hillman et al., 1999). Generally, rewards showed higher positive effects on performance among low performers as compared to high performers (Doran et al., 2008; Vaghela et al., 2008; Lindenauer et al., 2007; Glickman et al., 2007; Coleman et al., 2007). In addition, studies found no clear relationship between reward size and performance in systems that used pay-for-performance schemes. One study found a strong relationship between the reward size and the rate of adoption of pay-for-performance programs (de Brantes et al., 2009). Reward size explained 89 to 95% of variation in program participation in this study.

Other studies showed that pay-for-performance programs that provided rewards at the individual level (e.g. Coleman et al., 2007; de Brantes et al., 2009; Rosenthal et al., 2008; Fairbrother et al., 2001) and at the group level (Gilmore et al., 2007; Chung et al., 2003; Greene et al., 2004) both produced positive results. There was only one exception to these results. The study done by Young et al. (2007) found no significant effect of pay-for-performance programs on adherence to quality standards. Programs that were aimed at the hospital level produced less significant results (Pearson et al., 2008; Glickman et al., 2007; Rosenthal et al., 2005; Hillman et

al., 1999; Hillman et al., 1998; Morrow et al., 1995). This could be due to lower incentive payment per provider in hospitals.

Further, Van Herck et al. (2010) provided guidelines for the design of future pay-for-performance programs. Future programs should (1) select and define targets based on the baseline level of improvement, (2) use both process and outcome criteria as target measures, (3) involve stakeholders in the design and implementation of the program, (4) implement a uniform payment program among all the payers, (5) focus on both quality improvement and achievement, and (6) pay the rewards at the individual and/or team level (Van Herck et al., 2010).

Only a few studies examined the effect of pay-for-performance programs on referrals (An et al., 2008; Srirangalingam et al., 2006; Grady et al., 1997). In one study, Grady et al. (1997) evaluated the impact on referral rate of three different approaches designed to increase mammogram referrals for patients aged 50 years and older. Physicians were randomly assigned to three treatment conditions: (1) education-only condition, (2) education plus cue enhancement strategy condition, and (3) education plus cue enhancement plus feedback and reward condition (Grady et al., 1997). Physician education included the presentation of charts regarding the historical incidence of breast cancer and the provision of information on the increasing rate of breast cancer among older women, the high correlation between breast cancer and age, and the positive effect of physician encouragement on mammography use. Cue enhancement strategy consisted of using mammography chart stickers and spaces for three mammography referrals and completions. The reward was an amount of money determined by the percentage of referred patients in a specified period of time. For example, the physician would receive $50 for a 50% referral rate. The study included 61 practices in Dayton, Ohio, and Springfield, Massachusetts

over a three-year period resulting in a sample of 11,426 patients. The results of the study showed that chart sticker cueing significantly increased referrals, completions, and overall compliance in comparison with education alone. However, the researchers did not find any improvement from performance-feedback and reward above and beyond the use of sticker cueing. The authors speculate that the reward was too small and isolated to have had an effect on performance in these practices.

Srirangalingam et al's (2006) study focused on the implementation of a broad-based pay-for-performance program by the UK's National Health Service (NHS). Researchers analyzed how referral patterns for diabetes care changed after the introduction of the new program. The results showed no significant improvement in the total number of referrals to secondary care 6 months after the implementation of the new pay-for-performance program. But, it showed improvement in the quality of referrals. During the period of the program the thresholds for referral of diabetic patients significantly decreased. Patients referred after the contract was implemented had significantly lower glycated haemoglobin compared to those referred prior to the contract. This suggested that physicians were acting more vigorously upon poor glycaemic control after the implementation of the new reward system (Srirangalingam et al, 2006).

Finally, a recent study done by An et al. (2008) used a randomized trial to compare the effect of a pay-for-performance program on the referrals to tobacco quitline services. Tobacco quitline services provide evidence-based methods for stopping tobacco usage. The pay-for-performance program offered $5000 if 50 patients were referred to tobacco quitline services. The study was a randomized trial that compared usual care in 25 clinics with the pay-for-performance program in 25 clinics. The results showed that clinics under the pay-for-performance program

referred 11.4% of their patients. The rate was significantly higher than the 4.2% referral rate for usual care clinics (An et al., 2008).

4.6 Referral Appropriateness

Much evidence exists concerning inappropriate referrals to hospitals and specialists (Hutchison, 1993; Fertig et al., 1993; Helliwell & Wright, 1991; Sladden & Graham-Brown, 1989; Samanta & Roy, 1988). Policy makers often regard high referral rates as inappropriate (Coulter, 1998; Roland, 1992). Determining referral appropriateness is a complex issue and merely concentrating on high and low referral rates implies that the norm of referral rates should be somewhere close to the present average (Marinker et al., 1988), but the average referral rate does not provide information on the desired level of or the acceptable variability in referral rate (Wilkin et al., 1989; Marinker et al., 1988). In order to better understand the appropriateness of referrals, the reasons for referral should be taken into account.

O'Donnell et al. (2000) categorized the reasons for referral into three categories of investigation and/or diagnosis, treatment, and advice and reassurance for the patient and/or GP. When judging referral appropriateness, we should take account of the objectives of each of the above categories (O'Donell, 2000). In addition, an appropriate referral must be necessary for the individual patient, timely and effective in the management of the disease, and attempt to minimize overall costs (Coulter, 1998).

Most studies on referral appropriateness have asked GPs and/or specialists to review a series of referrals (Elwyn & Stott, 1994; Fertig et al., 1993; Emmanuel & Walter, 1989; Grace & Armstrong, 1987). In one study, 55% of hospital consultants across a range of specialties indicated that GPs could have done more before referring the patients (Grace & Armstrong,

1987). In Emmanuel & Walter's study, specialists felt that most of the referrals to hospitals were appropriate for hospital management after post-referral discussions between GPs and specialists were done (Emmanuel & Walter, 1989). In another study done by Fertig et al. (1993) in Cambridge, consultants reviewed 521 GP referrals and judged that only 9.6% of the referrals were inappropriate. In the same study, GPs reviewed 308 referral cases using referral guidelines and judged 15.9% of the cases to be inappropriate. In Elwyn and Stott's (1994) study, GPs reviewed a sample of referrals and found 34% of the referrals to be inappropriate. Most of these inappropriate referrals were due to lack of resources, lack of knowledge or required specialist skills and procedures.

Several studies have examined the contribution of inappropriate referrals to the variation in referral rates (Fertig et al., 1993; Coulter et al., 1990). Fertig et al. used specialist judgment to find the rate of inappropriate referrals. Except in orthopedic cases, they found 15% of the referrals to be inappropriate. In this study, Fertig et al. found that elimination of inappropriate referrals would have reduced the referral rates variation from 2.5- to 2.1-fold. Another interesting finding of this study was that if referral guidelines were used strictly, the absolute number of patient referrals would have increased. Therefore, in order to have a 100% appropriate referral rate, health care systems need to improve the effectiveness of their referral systems rather than to decrease the referral rate (Fertig et al., 1993).

Referral rates of GPs cannot reveal much about the appropriateness of their referrals. Indeed, GPs with an average referral rate may refer as inappropriately as those who have high or low referral rate (O'Donnel et al., 2000). Having a high, average, or low referral rate is not as important per se as the percentage of appropriate referrals made by the GP (O'Donnell, 2000).

Therefore, in order to improve referral effectiveness, inappropriate referrals should decrease where there is no benefit for patient referral and appropriate referrals should increase where there is a benefit for patient referral (O'Donnell, 2000).

There are several methods used for assessing referral appropriateness. It is important to note that the perspectives of the specialist, the general practitioner, and the patient on the referral appropriateness are probably different (Roland, 1992, Grace & Armstrong, 1987). From the medical perspective, referral appropriateness should be determined based on the views of specialists and GPs. As a result, joint groups of specialists and GPs have been established to develop guidelines for referral (Haines & Armstrong, 1992; Emmanuel & Walter, 1989).

An important factor in managing referral processes is effective communication between GPs and specialists regarding referrals. Coordinated communication between the GP and the specialist is defined as communicating timely, relevant patient information and the reason for the referral by GPs referring patients to specialists and communicating findings and recommendations by specialists back to the GPs (Williams et al., 1960). One study conducted in Canada examined the effectiveness of communication between GPs and specialists in Canadian health care (O'Malley & Reschovsky, 2011). This study showed that while 69.3% of GPs reported "always" or "most of the time" sending notification of a patient's history and reason for consultation to specialists, only 34.8% of specialists said they "always or "most of the time" received such notification. Similarly, 80.6% of specialists reported they "always" or "most of the time" send consultation results to the referring GPs, but only 62.2% of GPs said they received such information (O'Malley & Reschovsky, 2011). The results of this study show that there is still a lack of proper communication between GPs and specialists in the referral process.

Most of the studies done on inappropriate referrals have only focused on patients who have been inappropriately referred to specialists. However, when estimating the rate of inappropriate referrals, it is important to consider those who may inappropriately have not been referred to specialists also (Fertig et al., 1993). The design of our experiment that is explained in chapter 5 allows us to address both aspects of referral appropriatness including appropriate referrals and non-referrals.

4.7 Use of Hypothetical Cases in Referral Studies

Our experimental studies are based on the use of hypothetical cases in two areas of rheumatology and respirology. Hypothetical cases are frequently used to study physician decision making (Kankaanpää et al., 2012; Langley et al., 1997; Langley et al., 1991; Nightingale, 1987b). The results of one study done by Langley et al. (1991) showed that hypothetical patient case descriptions provided a useful tool for studying different factors affecting referral decisions. General practitioners responded appropriately to the information included in the hypothetical patient case descriptions (Langley et al., 1991).

One of the limitations of the use of hypothetical cases in studying medical decision making might be the fact that it is likely that physician's responses to hypothetical cases may not capture all the facets of the medical decision making environment (Jones, 1991; Rethans, 1991). However, the use of hypothetical cases will enable us to control for many environmental factors that may directly or indirectly affect decision making. Thus, this type of study will give useful insights regarding the effect of clearly defined factors on decision making within a well-defined context.

4.8 Conclusion

In this chapter we reviewed the factors explaining variation in general practitioners' referrals. Researchers categorized relevant factors to four categories of patient characteristics, practice characteristics, GP characteristics, and payment systems. In the next two chapters, we will describe the experiments that we conducted in order to examine the effect of two referral-related factors on referral pattern: GP-specialist relationship and payment schemes.

As was mentioned in the previous section, the use of hypothetical cases may be considered a limitation to the generalizability of our study results to the real world practices. However, physicians' decisions regarding hypothetical cases can give useful insight about their decision making in real contexts (Langley et al., 1991; Jones et al., 1990). In fact, creating a controlled setting is the main purpose of experimental design and analysis. To increase the generalizability of our results, we have incorporated some nonmedical factors into our cases (Langley et al., 1991; Rethans, 1991; Jones 1991). The purpose of the next two studies is to examine the effect of social relationships between GPs and specialists and payment schemes on referral rate and referral appropriateness. In order to test these hypotheses, we need to create a controlled environment to distinguish and examine the effect of these factors on referral patterns.

Chapter Five: Experimental Study on Social Relationships and Referral Patterns

5.1 Introduction

Task referral is an important issue in many operational systems such as health care systems. In health care, there has been a shift towards centralized referral systems in order to decrease waiting time and increase the flexibility of the processes. From a queuing theory perspective, a centralized referral system characterized by a central queue for all the referrals reduces average waiting time in the queue by mitigating the variability in the system (Dijk & Sluis, 2008; Rothkopf & Rech, 1987; Smith & Whitt, 1981). Therefore, the use of centralized or pooled referral systems has been recommended as a method to reduce patients' time to appointment by reducing the impact of variation in individual specialists' schedules on the time to appointment.

However, the existence of social relationships between general practitioners (GPs) and specialists in decentralized referral systems may decrease the total number of referrals by decreasing the rate of referrals. In this research, we have designed an experiment to investigate the effect of social relationships on the rates of referrals to specialists under a centralized referral system versus a decentralized referral system characterized by a close relationship between the general practitioners and the specialists.

Understanding different clinical and non-clinical factors affecting the medical referral process has always been of interest to researchers in a wide variety of areas including medicine, psychology, social psychology and operations management. As discussed in chapter 3, empirical

studies have confirmed the existence of systematic variation in rates of referral in general practice (McBride et al., 2010; O'Donnell, 2000).

Several studies have examined the effect of GP characteristics on referral decisions (e.g. De Marco et al., 1993; Reynolds et al., 1991). The GP-specialist relationship is one determinant of referral patterns in decentralized referral systems (Newton et al., 1991). In decentralized referral systems, GPs refer patients directly to the specialists whom they know, so that over time a professional relationship is developed between them. Knowing the specialist could influence a GP's referral decision making especially when the GP is uncertain whether the patient should be referred or not. Such close social relationships could motivate the GPs to be more careful about their referrals, as the GP could be mindful of not wasting the specialists' time or may feel guilty about referring patients who they might be able to manage themselves. Further, GPs might attempt to maintain their relationship and status before the specialists by not sending patients that are not urgent (Newton et al., 1991). In other words, social relationships can act as a mechanism towards lower referral rates by motivating GPs to pay more attention to their referral decisions. In one study, Langley et al. did not find a conclusive result on the impact of GP-specialist relationship on referral rate, although there was a significant variation among different groups of GPs about the effect of GP-specialist relationship on referral rates (Langley et al., 1997).

In this study, we have designed an experiment to examine the effect of social relationships between GPs and specialists on referral rates. We would like to examine the effect of social relationship on behavior between two experimental conditions in the context of referrals, which is an example of what is known in the literature as "gatekeeping". We will

compare referral rates under various experimental treatments to see if social relationships affect the referral rate.

5.2 Hypotheses

We hypothesize that in the presence of social relationships, which are characteristic of decentralized referral systems, the number of referrals will be significantly smaller than in the absence of such factors.

Hypothesis 1. Doctors in a decentralized referral system characterized by a close relationship will refer significantly fewer patients to specialists in comparison with the doctors in a centralized referral system.

Following hypothesis 1, predicting that doctors in a close-relationship condition refer fewer patients to specialists in comparison with doctors in a centralized referral system, we hypothesize that doctors who have a close relationship with the specialists would give lower urgency ratings to their patients in order to justify their decision not to refer the patients.

Hypothesis 2. Doctors in a decentralized referral system will give a significantly lower average urgency rating to their patients in comparison with doctors in a centralized referral system.

5.3 Experimental Design

We designed an experiment to examine the effect of a change in the relationship between the GP and the specialist on referral rate and referral urgency. We hypothesized that participants in the close relationship condition (decentralized referral system) would refer significantly fewer patients to specialists in comparison with participants in the centralized referral system. To test this hypothesis, we conducted an online survey with family practice residents as participants. 55

family practice residents participated in the survey and were randomly assigned to a 2x2 experimental design based on the type of the relationship between the GP and the specialist (2 types: close relationship (decentralized system) and no relationship (centralized referral system)) and test type (2 types: rheumatology and respirology).

The study was done online through Qualtrics software. Qualtrics is a global data collection and analysis provider for research and data collection in different areas including market research, voice of customer, employee performance, and academic research. It provides a platform for data collection and analysis (http://www.qualtrics.com/blog/about-qualtrics/). The University of Calgary is one of the universities that uses the Qualtrics platform for quantitative and qualitative research purposes.

The design of the experiment was a mixed-method matched and random assignment. The software randomly assigned participants to different treatment conditions. Participants were told that the purpose of the study was to understand the factors affecting priority setting in referral process from GPs to specialists.

In order to manipulate the type of relationship between the participant as the GP and the specialist in each treatment, we provided a description of the relationship between the participant and the specialist to whom the patient may be referred. The relationship between the GP and the specialist was one of two types: (1) close relationship, which is characteristic of decentralized referral systems and (2) no relationship which is characteristic of anonymous centralized referral systems. After reading the description about the relationship between the GP and the specialist, participants were presented with 12 patient case descriptions. The case descriptions were written based on real patients who had visited the specialists in the two areas of rheumatology and

respirology. However, they were completely anonymous and were given no labels in order to avoid any probable influences on the respondents. They were either in the rheumatology or respirology area of specialty. The rheumatology cases were developed by the Western Canada Waiting List group (WCWL) and provided to us by Dr. Tom Noseworthy, Professor of Health Policy and Management in the Department of Community Health Sciences and Institute for Public Health, University of Calgary. The Respirology cases were developed and provided to us by Dr. Sachin Pendharkar, Assistant Professor in the University of Calgary. For each patient case description, participants answered three questions: (1) whether they would refer the patient to the specialist, (2) the urgency or relative priority of the patient for referral on a scale from 0 (Not Urgent) to 10 (Very Urgent), and (3) a brief explanation about the reasons for their referral or non-referral decision. The 0-10 scale was a discrete version of the Visual Analogue Scale (VAS) used as the Priority-Referral Score (PRS) by the Western Canada Waiting List (WCWL) group for rating the relative urgency of referrals by general practitioners (De Coster et al., 2007). Appendix B presents all parts of the experiment including the survey description, preliminary questionnaire, case test, and after-the-test questionnaire.

5.4 Variables

5.4.1 Dependent Variables

Table 5.1 provides a list of all the variables that were measured and included in this experiment. The study had two main dependent variables: the total number of referrals from GP to specialist and the average urgency rating for each participant. The case test had 12 patient case descriptions. For each patient case description, the participants answered whether they would refer the patient to a specialist (yes/no question) and how urgent it was for the patient to visit the

specialist on a 0-10 scale. Referral rate was calculated by counting the total number of positive responses by the participant on all the patient case descriptions. The average urgency rating was calculated by finding the average of the urgency ratings on the 12 patient case descriptions.

5.4.2 Independent Variables

The study has two main independent variables: relationship type and test type that are operationalized using dummy variable coding. Social relationship had two levels of no relationship (centralized referral system) and close relationship (decentralized referral system). Social relationship is set to -0.5 for centralized referral system conditions and is set to 0.5 for close relationship conditions. Test type has two types of rheumatology and respirology. Test type variable (Rheum) is set to 0.5 for the rheumatology condition and is set to -0.5 for the respirology condition. In order to test our research hypotheses, we used contrast coding (-0.5/0.5) coding instead of dummy coding (0/1) coding for our categorical independent variables. Contrast coding is the best method for testing hypotheses related to main effects and interaction in multiple regression models (Cohen & Cohen, 1985).

Dummy variables are similar to uncentered quantitative variables. The use of dummy variable coding instead of contrast coding may produce interpretable results. The significance of the interaction term is not influenced by the use of dummy variable coding scheme. However, the absence of centering which is characteristic of contrast coding makes the main effects of the variables difficult to interpret (Cohen & Cohen, 1985).

Table 5.1 Variables

Variable Name	Variable	Measure	Variable Type
Relationship (Close)	Independent	Type of relationship between the participant as the GP with the specialist	Nominal with two levels : 0.5 = Close relationship; -0.5 = Centralized referral system
Test Type (Rheum)	Independent	Area of specialty/Test type	Nominal with two levels : 0.5 = Rheumatology; -0.5 = Respirology
Referral self-confidence (RefConf)	Independent	The confidence in referral decision making ability	Scale (1-7)
Referral Rate (RefRate)	Dependent	Number of referrals	Scale (0-12)
Urgency (AvgUrgency)	Dependent	Average of urgency ratings on the 12 patient case descriptions	Scale (1-10)
Mood	Dependent	Two items selected from PANAS (Positive and Negative Affect Scale) mood scale	Scale (1-7)
Knowledge (Know)	Covariate	Level of knowledge in the areas of rheumatology or respirology depending on the test type	Scale (1-7)
Age	Covariate	Participant's age	Ordinal based on age group
Gender	Covariate	Participant's gender	Nominal with two levels: 1 = Female; 0 = Male
Year of program (Year)	Covariate	Participant's year of residency program	Scale
Location of program (Rural)	Covariate	Participant's location of residency program	Nominal with two levels: 1 = Rural; 0 = Urban
Primary Language (English First)	Covariate	Participant's primary language	Nominal with two levels: 1 = English; 0 = Other
Risk	Covariate	Selected items from Holtgrave et al.'s (1990) multidimensional risk scale	Scale (1-7)
Loss Aversion (LossAverse)	Covariate	One item used in Nightingale's studies (1987a,b)	Nominal with two levels: 0 = Risk seeking; 1 = Loss Averse

5.4.3 Control Variables

In addition to primary measures of interest, we collected data on a set of other variables in order to control for the effect of these variables on both dependent variables of referral rate and referral urgency. Based on the referral literature, GP characteristics can have a significant

effect on referral decision making. First of all, the literature shows that a GP's knowledge and confidence towards a specific area of specialty may either decrease (Newton et al., 1991) or increase (Reynolds et al., 1991; Morrell et al., 1971) referrals in that specific area of specialty. In order to measure knowledge level in each area of specialty, we asked participants to indicate their level of knowledge in both areas of rheumatology and respirology prior to the case test on a scale from 0 (not at all knowledgeable) to 7 (very knowledgeable). In order to measure confidence, we asked participants to rate their confidence level using one item: ("Please rate your confidence in your ability to determine whether a patient of yours should be referred to a specialist in the two areas of rheumatology and respirology"; 0 = not at all confident; 7 = very confident). Participants' knowledge and confidence were measured prior to the case test in the preliminary questionnaire. Appendix B.2 presents the preliminary questionnaire.

Second, studies have found that a GP's risk seeking preference and degree of loss aversion can significantly impact different clinical factors such as laboratory usage (Holtgrave et al., 1990) and referral to hospital (Nightingale et al., 1988; Nightingale, 1987a, 1987b). In order to measure participants' risk attitude, we used a multidimensional scale because several studies have shown that risk attitude is multidimensional (Jackson et al., 1978; Weber, 1988). Similarly, risk attitude in medical decision making must consider different aspects of the individual's risk seeking preferences (Holtgrave et al., 1990). Therefore, using a multidimensional scale, we measured risk in a number of different domains including medical, financial, social, and physical domains. Tables 5.2 and 5.3 provide the scale and binary items of the risk attitude measure used in our study.

The risk attitude measurement instrument consisted of an 8-item measure, which participants voluntarily completed after they finished the case test. The survey instrument was placed after the case test in order to avoid any confounding effect of individual's performance in the case test with risk measurement. The risk attitude instrument was a reduced version of Holtgrave et al.'s (1990) multidimensional risk scale. It comprised items both from Nightingale's risk measurement in medical domain (Nightingale et al., 1988; Nightingale, 1987a, 1987b) and Jackson's standardized risk measurement method (Jackson et al., 1978) that covered risk attitudes in financial, physical, and social domains. In order to keep the survey length short, we decided to use an abbreviated version of the risk attitude instrument used by Holtgrave et al. in this study.

Table 5.2 Seven-point Likert scale risk items

Seven-point Likert scale items 1(Not at all true)-7 (Very true)
I am outspoken. (Feelings)
I am concerned about getting hurt. (Danger)
I enjoy an element of physical danger. (Danger)
I am careful about not hurting people's feelings. (Feelings)
I am generally risk-averse when dealing with patients (RiskPatient).

Table 5.3 Binary choice risk items

Binary choice items
I would not speak out in public on unpopular issues unless I were fairly certain of my position. (Public) True/False
If asked to choose between the two following therapies for your patient, which one would you select? (Gain) ☐ Therapy A, which **definitely** gives the patient five more years of life than the average longevity ☐ Therapy B, which gives a 50:50 chance of zero or ten years of additional life for your patient your patient
If asked to choose between the two following therapies for your patient, which one would you select?(Loss) ☐ Therapy A, which the patient definitely lives five years less than the average person ☐ Therapy B, which gives a 50:50 chance of losing zero or ten years of additional life for the patient

Third, we controlled for geographic location by asking the participants' location of their residency program (urban versus rural). Based on the referral literature, there is regional variation in the management of medical problems (Pilote et al., 1995; Notzon et al., 1987) and in the use of medical and surgical services such as services that are made available through the referral process (Birkmeyer et al., 1998; Wilkin & Smith, 1987a, 1987b; Chassin et al., 1986; Wennberg, 1986; Boyle, 1985; McPherson et al., 1985). A GP's geographic location is an important indicator of referral rates (Langley et al., 1997, 1991; Jones, 1987); rural GPs have significantly lower referral rates in comparison with urban GPs (Langley et al., 1997, 1991; Madeley et al., 1990). Finally, we collected participants' demographic data including age group, gender, ethnicity, and whether English was their first language.

5.4.4 Other Dependent Variables

Another dependent variable measured in our study was the participant's mood. We measured mood to test the likelihood of the impact of the social relationship between the

participant as the GP and the specialist on participant's mood. There is a large body of literature on the relationship between mood and social relationships (McIntyre et al., 1991; McIntyre et al., 1990). For example, some experimental studies showed that positive versus negative mood was associated with greater interest in social relationships (Cunningham, 1988a, 1988b). Other studies showed that spontaneous or arranged social interactions increased positive affect in comparison with a neutral control condition (McIntyre et al., 1991; McIntyre et al., 1990). We hypothesized that our manipulation would not affect participants' mood. Therefore, any differences in the referral rate and urgency ratings could be attributed to the difference in the social relationship between the participant as the GP and the specialist and are not related to the participants' mood variation in different treatment conditions.

In order to keep the survey length short, we used a parsimonious 2-item mood scale selected from the PANAS scale (PANAS; Watson et al., 1988) to measure participants' mood during the experiment. Mood measure was placed immediately after the case test section. On a 7-point scale, we asked participants to indicate how much they agreed with each of the following two sentences: (1) I felt cheerful while I was completing the test and (2) I felt upset while I was completing the test (1 = not at all true; 7 = very true).

The results of a univariate analysis of variance (ANOVA) on the effect of relationship type on mood showed that our manipulation for the relationship type did not have a significant effect on mood ($F = 0.013$; $p = 0.911$). Therefore, the changes in referral rates and urgency ratings could not be attributed to changes in the participants' mood.

5.5 Participants and Procedure

The online survey instrument was emailed to family practice residents through the Departments of Family Medicine in seven Canadian universities including University of Calgary, University of British Columbia, University of McGill, University of Ottawa, University of Manitoba, University of Saskatchewan, and Memorial University of Newfoundland. The approximate sample of residents who received the email was 900 people, among which 55 family practice residents participated in this study. The survey was sent to residents four times in October, and December, 2012, and January and February, 2013. Approximately 40% of the sample was obtained in the third round in January 11, 2013. This can be due to the fact that the survey was sent right after the Christmas holidays when participants were back from holiday breaks, so they were fresh and less busy in comparison with other rounds.

The participants were randomly assigned to the conditions of a 2 (social relationship: close relationship vs. centralized referral system) x 2 (test type: rheumatology vs. respirology) between subjects factorial design. Participants first read a survey description which consisted of information on the survey purpose, its content, the approximate length of time it took to complete the survey, and the possibility of taking part in a draw for an Apple iPad 3 (value of $550) after completing the survey. The purpose of the survey as explained to participants was “to understand the factors affecting priority setting in the patient referrals from general practitioners to specialists”. Appendix B.1 provides the complete online survey description.

In section 1, participants answered a preliminary questionnaire including questions on their level of knowledge and confidence in answering questions in two areas of rheumatology and respirology. Appendix B.2 presents the preliminary questionnaire.

Section 2 included the case test, which consisted of 12 patient case descriptions developed based on real patients who were referred by general practitioners to specialists in rheumatology and respirology. In order to manipulate the relationship between the participant as the general practitioner and the specialist to whom they might refer their patients, we provided a written description of the relationship between the participant and the specialist. In close relationship conditions, participants read an explanation about a specialist in their town to whom they had developed a close professional relationship after having referred many of their patients to this specialist over the past few years. In the centralized referral system conditions, on the other hand, participants were provided with an explanation of a centralized referral system to which they would refer their patients. Table 5.4 presents the manipulation for social relationship in each of the four treatment conditions.

Table 5.4 Social relationship manipulation

a) Rheumatology/close relationship condition The following case describes one of your patients who you may consider referring to a rheumatologist. The rheumatologist to whom you may refer the patient is Dr. Lee. Having referred many of your patients to Dr. Lee over the past few years, you consider yourself to have a close professional relationship with Dr. Lee. You trust and feel comfortable with Dr. Lee as a specialist. Please read the case and answer the questions below.
b) Respirology/close relationship condition The following case describes one of your patients who you may consider referring to a respirologist. The respirologist to whom you may refer the patient is Dr. Lee. Having referred many of your patients to Dr. Lee over the past few years, you consider yourself to have a close professional relationship with Dr. Lee. You trust and feel comfortable with Dr. Lee as a specialist. Please read the case and answer the questions below.
c) *Rheumatology/no relationship condition* The following case describes one of your patients who you may consider referring to a rheumatologist. The referral system to which you refer the patient is a **centralized referral system** in which the patients are referred to different rheumatologists through a centralized system that allocates the patient to the rheumatologist that has the next available time slot. Please read the case and answer the questions below.
d) Respirology/ no relationship condition The following case describes one of your patients who you may consider referring to a respirologist. The referral system to which you refer the patient is a **centralized referral system** in which the patients are referred to different respirologists through a centralized system that allocates the patient to the respirologist that has the next available time slot. Please read the case and answer the questions below.

Considering the social relationship description, participants were then asked to read each patient case description and decide whether they would refer the patient to this specialist or not. Appendix B.3 presents a shortened version of the case test section of the study.

Finally, in section 3 participants answered an after-the-test questionnaire that consisted of questions on mood (PANAS, Watson et al., 1988), risk attitude, geographic location of residency

program, and demographic data. Data on factors including age, gender, ethnicity, and primary language was also collected. Appendix B.4 presents the after-the-test questionnaire.

5.6 Main Statistical Analyses and Results

Descriptive statistics of the data showed that 61% of the participants were female and 39% were male. There was not a significant difference in the female/male mix in each of the four treatment conditions. In addition, 64% of the participants were working in the urban areas while 36% were working in the rural areas. There was not a significant difference in the number of rural versus urban participants in each of the four treatment conditions. Participants were all either in the first or second year of their residency program. Only one participant had already finished his/her residency program. Year of residency for this participant was entered as 3 in our analysis. Table 5.5 provides the descriptive statistics for all the dependent, independent, and control variables in total and in each treatment condition separately.

In order to be able to test the difference in referral behavior between the centralized referral system versus the close relationship condition among high-confidence and low-confidence individuals, we converted the continuous variable "Confidence" to a categorical variable called "RefConf". Categorizing the confidence variable allowed us to test the effect of relationship type on referral rate for high-confidence and low-confidence participants separately. The RefConf value for those whose confidence level was less than 5 was coded as -0.5, while the RefConf value for those whose confidence level was more than or equal to 5 was coded as 0.5.

Table 5.5 Descriptive statistics

Descriptive Statistics	All Treatments			RheumClose			RheumCentral			RespClose			RespCentral		
Variable	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**
RefRate	6.05	2.60	55	7.87	2.19	16	7.75	2.45	12	4.09	1.51	11	4.31	1.30	16
AvgUrgency	4.07	1.52	55	4.86	1.46	16	5.05	1.35	12	3.39	0.96	11	3.03	1.16	16
Mood	4.02	1.03	51	3.93	0.96	14	3.82	0.78	11	4.2	0.79	10	4.13	1.38	16
Rheum	0.01	0.50	55	0.5	0	16	0.5	0	12	-0.5	0	11	-0.5	0	16
Close	-0.01	0.50	55	0.5	0	16	-0.5	0	12	0.5	0	11	-0.5	0	16
RefConf	0.04	0.50	54	0.1	0.51	15	0.33	0.39	12	-0.05	0.52	11	-0.19	0.48	16
Know	3.58	0.94	55	3.44	1.09	16	3.5	0.80	12	3.82	1.17	11	3.62	0.72	16
Risk	2.33	0.68	55	2.28	0.72	16	2.49	0.70	12	2.22	0.75	11	2.35	0.61	16
LossAverse	0.66	0.48	50	0.75	0.45	16	0.64	0.50	11	0.6	0.52	10	0.62	0.51	13
Rural	0.36	0.48	45	0.38	0.51	13	0.1	0.32	10	0.5	0.53	10	0.42	0.51	12
Gender	0.61	0.49	54	0.73	0.46	15	0.42	0.51	12	0.64	0.50	11	0.62	0.5	16
Age	1.31	0.47	55	1.31	0.48	16	1.25	0.45	12	1.18	0.40	11	1.44	0.51	16
Year	1.40	0.53	52	1.29	0.47	14	1.42	0.51	12	1.64	0.67	11	1.33	0.49	15
EnglishFirst	0.89	0.31	55	0.94	0.25	16	0.75	0.45	12	0.91	0.30	11	0.94	0.25	16
Valid N			40			9			9			10			12

In order to test our first hypothesis regarding the effect of social relationships and other related independent variables on referral rate, we ran a multiple regression analysis. In order to test the interactions between independent variables of interest, we included all the two-way and three-way interactions between the type of relationship, test type and confidence in the model.

In order to avoid multicollinearity, we added variables in a step-by-step format: In the first step, we entered the control variables only. In the second step, we entered the three independent variables: Rheum, Close, and RefConf. Table 5.6 shows the result of the multiple regression analysis in the step-by-step format. Model 1 includes only the control variables. Model 2 includes the control variables and the independent variables. Model 3 includes control variables, independent variables, and the two-way and three-way interactions. Model 4 and 5 are the follow-up tests for the effect of different referral systems on referral rate among individuals with high and low level of referral self-confidence. We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. In this table, * indicates that the variable is significant at 0.1, ** indicates that the variable is significant at 0.05, and *** indicates that the variable is significant at 0.01.

Table 5.6 Multiple regression: dependent variable: referral rate

	Dependent Variable: Referral Rate				
	Main Test			Follow-up Tests	
				High RefConf	Low RefConf
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Constant	7.345 (0.002)	4.704 (0.001)	4.351*** (0.001)	3.739 (0.035)	5.634 (0.017)
Knowledge (1-7)	-0.121 -0.048 (0.757)	0.425* 0.163 (0.082)	0.535** 0.205 (0.022)	0.608** 0.218 (0.038)	0.024 0.009 (0.962)
Loss Aversion (0/1)	1.270 0.234 (0.137)	0.996** 0.184 (0.035)	1.253** 0.231 (0.006)	1.594** 0.269 (0.024)	0.383 0.076 (0.668)
Risk	-0.304 -0.086 (0.603)	-0.212 -0.059 (0.503)	-0.36 -0.101 (0.227)	-0.503 -0.148 (0.158)	0.418 0.105 (0.549)
Rural (0/1)	-1.743 -0.327 (0.040)	-0.571 -0.107 (0.228)	-0.510 -0.096 (0.246)	-0.293 -0.056 (0.593)	-0.813 -0.144 (0.371)
Gender (0/1)	-0.235 -0.044 (0.788)	0.440 0.082 (0.360)	0.547 0.103 (0.234)	0.459 0.087 (0.450)	0.052 0.009 (0.957)
Rheumatology (Rheum) (-0.5/0.5)		4.507*** 0.881 (0.000)	4.369*** 0.854 (0.000)	3.981*** 0.759 (0.000)	4.524*** 0.876 (0.000)
Relationship Type (Close) (-0.5/0.5)		-0.396 -0.077 (0.362)	-0.184 -0.036 (0.654)	-1.368** -0.263 (0.020)	0.223 0.045 (0.767)
Referral Confidence (RefConf) (-0.5/0.5)		-0.867* -0.169 (0.080)	-0.794* -0.155 (0.095)		
Rheum*Close			0.798 0.078 (0.331)		
Close*RefConf			-2.171** -0.212 (0.015)		
Rheum*RefConf			-0.763 -0.073 (0.38)		
Rheum*Close*RefConf			-2.722 -0.132 (0.102)		
N	42	41	41	22	18
R Squared (%)	16.8	79.4	84.8	87.3	81.7
Adjusted R Squared	5.6	73.6	78.5	81.3	70
F Statistics	1.499	13.675***	13.453***	14.679***	7.006***
P value	0.214	0.000	0.000	0.000	0.002

We did not find any significant effect of relationship type on referral rate to support our first hypothesis. Interestingly, we found a significant interaction between relationship type and referral self-confidence in the main test (β = -2.239; $p < 0.025$).

Other variables that had a significant effect on referral rate consist of Rheumatology, Knowledge, and Loss Aversion: First, there was a higher referral rate for rheumatology cases in comparison with respirology ones. Residents referred rheumatology cases significantly more than respirology ones (β = 4.346; $p < 0.001$). Second, knowledge had a significant positive effect on referral rate (β = 0.586; $p < 0.025$). Interestingly, residents who had more knowledge referred more patients to specialists in comparison with those who had less knowledge of the area. This result is in line with the previous empirical studies on the effect of knowledge and interest on referral rate (Newton et al., 1991; Reynolds et al., 1991; Evans & McBride, 1968). Based on these studies, physicians who had higher knowledge and interest in a specific area referred more patients to specialists in these areas.

Third, loss aversion had a positive relationship with referral rate. Those who were loss averse referred significantly more patients (β = 1.146; $p < 0.025$). This result was in line with the results of Nightingale's studies on the effect of loss aversion on laboratory use and referral to hospitals (Nightingale, 1988; 1987 a; 1987 b). In a series of studies, Nightingale had found that doctors who were loss averse ordered more laboratory tests (Nightingale, 1987 a; 1987 b) and referred more patients to hospitals (Nightingale, 1988).

In order to follow up on the interaction between the Close and Confidence variables to test our hypotheses regadring the effect of relationship type on referral rate among low confidence and high confidence doctors, we ran two separate multiple regressions to examine the

effect of relationship type on referral rate among participants with low referral self-confidence (RefConf = -0.5) and those with high referral self-confidence (RefConf = 0.5) (Hair et al., 2010).

The results of the multiple regression for high-confidence individuals showed that participants in the close relationship treatments referred significantly fewer patients in comparison with participants in the centralized referral system: Close was significant among high-confidence individuals (β = -1.385, $p < 0.025$). In addition to the main variable Close, other control variables including Know (β =0.693, $p < 0.05$) and LossAverse (β =1.426, $p < 0.05$) were both significant and had a positive effect on referral rate among high-confidence individuals. Model 4 of Table 5.6 shows the result of the multiple regression analysis on the effect of relationship type on referral rate among high-confidence individuals.

The second multiple regression for low confidence individuals showed no significant relationship between Close and RefRate. Model 5 of Table 5.6 shows the result of the multiple regression analysis for low-confidence individuals. Figure 5.1 illustrates the interaction between relationship type and confidence and its effect on referral rate.

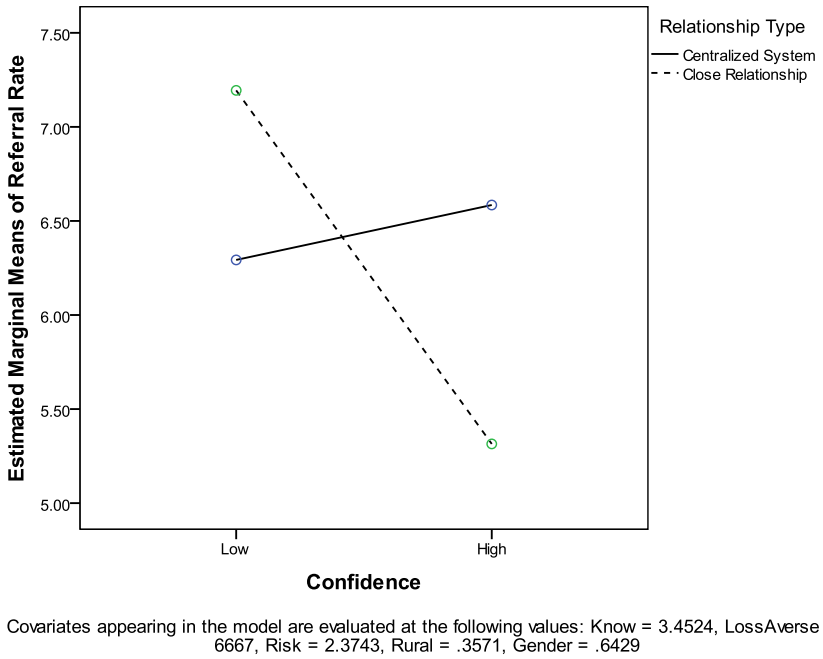


Figure 5.1 Relationship x confidence interaction and referral rate

In order to test the second hypothesis on the effect of the type of relationship on referral urgency, we ran multiple regressions in a step-by-step format as we did for referral rate. The results of the analysis did not show a significant effect of the type of relationship on average referral urgency (β = 0.229, p = 0.517). However, there was a significant interaction between relationship type and confidence (β = -1.432, p = 0.057). The follow-up tests did not indicate any significant difference in referral urgency between centralized system and close relationship among high-confidence and low-confidence individuals. Table 5.7 illustrates the multiple regressions on the effect of different independent variables on referral urgency. In model 1, the effect of control variables on referral urgency is tested. In model 2, independent variables are added, and in model 3, all two-way and three-way interactions are added to the regression model.

Models 4 and 5 consist of the follow-up tests for the effect of relationship type on referral urgency among low-confidence and high-confidence individuals.

Figure 5.2 depicts the effect of relationship type on referral urgency among low-confidence and high-confidence individuals. While the results of the multiple regression analysis are not statistically significant, the direction of the relationship between relationship type and referral urgency for high-confidence residents is in the direction that we predicted, i.e, physicians who had high confidence gave lower referral urgency to their patients in the close relationship condition in comparison with the centralized referral system.

Table 5.7 Multiple regression_dependent variable: referral urgency

	Dependent Variable: Referral Urgency				
	Main Test			Follow-up Tests	
				High RefConf	Low RefConf
	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	4.898 (0.000)	4.159 (0.000)	4.068 (0.000)	3.956 (0.012)	4.092 (0.029)
Knowledge (1-7)	0.169 0.116 (0.461)	0.306 0.206 (0.106)	0.307 0.206 (0.110)	0.403* 0.247 (0.099)	0.032 0.023 (0.939)
Loss Aversion (0/1)	-0.073 -0.023 (0.883)	-0.274 -0.089 (0.441)	-0.257 -0.083 (0.479)	-0.302 -0.087 (0.590)	-0.312 -0.123 (0.668)
Risk	-0.187 -0.092 (0.584)	-0.157 -0.077 (0.522)	-0.236 -0.115 (0.344)	-0.291 -0.147 (0.333)	0.152 0.076 (0.787)
Rural (0/1)	-0.608 -0.200 (0.212)	-0.158 -0.052 (0.666)	-0.065 -0.021 (0.859)	0.182 0.060 (0.698)	-0.485 -0.171 (0.508)
Gender (0/1)	-0.995* -0.327 (0.057)	-0.698* -0.229 (0.067)	-0.497 -0.163 (0.200)	-0.524 -0.171 (0.318)	-0.539 -0.190 (0.499)
Rheumatology (Rheum) (0/1)		2.072*** 0.709 (0.000)	1.894*** 0.648 (0.000)	2.364*** 0.772 (0.000)	1.369* 0.529 (0.076)
Relationship Type (Close) (0/1)		0.046 0.016 (0.891)	0.230 0.079 (0.509)	-0.582 -0.192 (0.217)	0.715 0.286 (0.257)
Referral Confidence (RefConf)		0.180 0.061 (0.632)	0.353 0.120 (0.370)		
Rheum*Close			0.332 0.057 (0.629)		
Close*RefConf			-1.435** -0.245 (0.051)		
Rheum*RefConf			0.726 0.122 (0.322)		
Rheum*Close*RefConf			-1.573 -0.133 (0.257)		
R Squared (%)	13	61	66.9	72.4	51.6
Adjusted R Squared	1.3	51.6	53.2	52.5	20.8
F Statistics	1.107	6.459***	4.889***	5.617***	1.676
P value	0.373	0.000	0.000	0.002	0.213
N	42	41	41	22	18

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

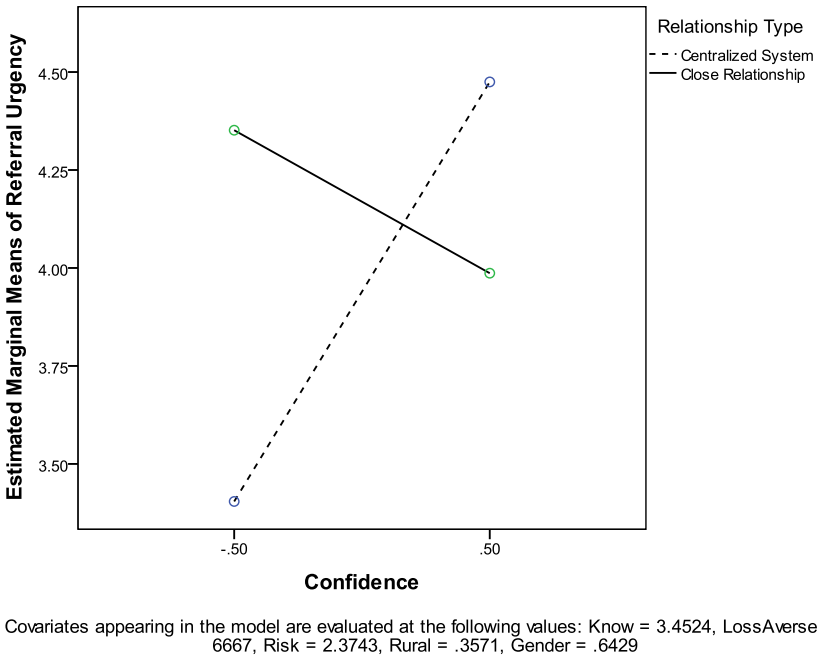


Figure 5.2 Relationship x confidence interaction and referral urgency

5.7 Conclusion

In this study, we examined the effect of social relationships among GPs and specialists on referral decision making. Our hypothesis on the effect of close relationship between GPs and specialists on referral rate was supported among residents who had high confidence in their referral decision making ability: close relationships between GPs and specialists decreased referral rate among family practice residents who perceived themselves to be able to make appropriate referral decisions. The results of our study indicate that high-confidence doctors are able to act in a "relationship-preserving" fashion by reducing their referrals while the low-confidence doctors cannot act in this fashion because they lack the confidence to deal with the patients themselves.

Our study sheds light on one of the socio-psychological factors influencing referral behavior, namely GP-specialist relationship. This result has important real-world implications. In recent years, there has been a shift from decentralized referral systems where GPs individually refer patients to specialists towards centralized referral systems which automatically allocate patients to the next available specialist. Based on queuing theory, centralized referral systems are more efficient and can facilitate the referral process by creating a central queue rather than multiple single queues for patients that get referred from different GPs to specialists (Rothkopf & Rech, 1987; Smith & White, 1981). However, social factors such as long-term social relationships that are developed between GPs and specialists in decentralized referral systems can safeguard against high referral rates: as was shown in this study, in decentralized referral systems, doctors who had high confidence in their referral decision making referred fewer patients than they did in centralized systems.

In addition to our main hypothesis, our study provided evidence on the effect of other personal and psychological factors on referral rate. In line with previous studies, our study showed that loss-aversion and knowledge both had a positive effect on referral rate. In addition, similar to previous studies, we did not find any significant relationship between risk seeking preference and referral rate.

However, our study has some limitations: First of all, our sample included family practice residents who do not have the knowledge and experience that primary care physicians will gain regarding the referral process over the years of their practice. In later studies, we may repeat this experiment with primary care physicians to gain a better understanding of the effect of social relationships on referral rate. Second, our study was based on the use of hypothetical cases in an

experimental setting rather than real-world patients. Conducting similar studies in real-world settings can give a better understanding of the nature of the referral process.

Chapter Six: Experimental Study on Payment Schemes and Referral Patterns

6.1 Introduction

Based on economic theories, payment systems can influence performance (Conrad & Christianson, 2004; Bull et al., 1991; Donaldson & Gerard, 1984; Evans, 1974). In health care systmes, payment schemes are often reformed as an attempt to enhance the performance of health care professionals (Levaggi & Rochaix, 2003). Empirical evidence shows that different types of payment systems could influence physician performance (Devlin & Sarma, 2008; Gosden et al., 2001; Gosden et al., 1999; Hughes & Yule, 1992; Goldmann, 1952). Health care literature provides evidence on the impact of different payment schemes on physicians' performance (McGuire, 2000; Scott & Hall, 1995; Stearns et al., 1992). Empirical evidence indicates that different payment methods could have different impacts on GP referral behavior (Hutchison et al., 2001).

Two main methods of remunerating physicians are fundholding and pay-for-performance. In the following sections, we will define each payment method's characteristics and examine its effects on GP referral behavior. Our literature review showed that previous studies had not tested the effect of different payment methods on physicians' referral behavior in a controlled experimental setting. Therefore, we designed an experiment to examine the effect of fundholding and pay-for-performance on referral rate and referral appropriateness in general practice.

6.2 Fundholding and Referral Behavior

Fundholding policy is considered one of the major types of payment reforms in health care systems. Under the fundholding policy, GPs receive budgets for purchasing elective

(planned) care from hospitals and/or other health care providers (Brereton & Vasoodaven, 2010). Fundholding policy was developed to provide the necessary incentives for GPs to manage their prescription and referral patterns more effectively and efficiently and to provide health care services that better reflect the requirements of their local population (Dixon & Gelennerster, 1995).

As was explained in the literature review chapter, empirical studies on the effect of fundholding on referral provide mixed results. Some studies found that fundholding decreased cost of service. For example, Surender et al. compared the referral rate of fundholder GPs and non-fundholder ones before and after the introduction of the fundholding policy. The results of this study showed that the referral rates of non-fundholders increased by 26.6 percent after three years while the referral rate of fundholders had increased by only 7.5 percent within the three-year period of the study (Surender et al, 1995). Dusheiko et al. (2006)'s study found that referral rates for chargeable elective admissions increased for ex-fundholders after the abolishment of the scheme (Dusheiko et al., 2006). This result can be due to the fact that GPs have tried to maximize their savings by decreasing their referral rates (Dusheiko et al., 2006).

Contrary to the results of the above-mentioned studies, Coulter and Bradlow (1993) found no significant difference in referral rates of fundholders and non-fundholders. They compared referral rates from 10 “first wave” fundholders with 6 non-fundholding practices during the preparation year (phase 1) and one year after the introduction of the fundholding scheme (phase 2). The fundholding budget for the first year was based on the referral rate in the preparatory year. After the initiation of fundholding reforms, both groups of practices increased their referral rates. The fundholding practices increased their referrals from 107.3 per 1000

patients per annum to 111.4 and the non-fundholders from 95.0 to 112. There was no significant difference in referral rate of fundholders and non-fundholders in this study (Coulter & Bradlow, 1993). In another study, Kammerling & Kinnear (1996) evaluated the effect of fundholding on outpatient referrals to orthopedic clinics one year before becoming a fundholder and two years after that. In the year before fundholding, the practices in the fundholding group referred fewer patients than the control group although the difference was not significant. However, two years after the fundholding program, fundholders' referral rate increased by 13% while non-fundholders referral rate increased by 32%. Therefore, as the values show, fundholders increased their referrals less than non-fundolders two years following the prgram (Kammerling & Kinnear, 1996). However, based on Akbari et al.'s review (2008) the Kammerling and Kinnear's study suffers from a unit of analysis error and its authors did not report the statistical significance of their findings (Akbari et al., 2008).

Based on the fundholding literature, it can be concluded that the results of studies on the effect of fundholding on referral are inconclusive and not generalizable since these studies have tended to be small scale and to lack adequate controls (Dusheiko et al., 2006). Therefore, further analysis in a controlled setting is still needed in order to get a better understanding of the effect of fundholding on referrals.

In this study, we examine the effect of fundholding policy on referral behavior in an experimental setting. Based on the literature, fundholding provides financial incentives for GPs to manage their referral budgets more efficiently and utilize the budget savings for other health care services based on their clinical judgment (Dixon & Gelennerster, 1995). Therefore, we hypothesize that:

Hypothesis 1. Medical doctors who work under the fundholding policy will refer significantly fewer patients to specialists in comparison with those who work under the fixed payment policy.

Hypothesis 2. Medical doctors who work under the fundholding policy will be biased towards lower urgency ratings in comparison with those who work under the fixed payment treatment.

6.3 Pay-for-Performance and Referral Behavior

Pay-for-performance is another payment method that has often been applied to health care systems in recent years. It directly relates a proportion of the remuneration of the physicians and health care providers to the achieved results on quality factors of their performance such as the improvement of the referral process (An et al., 2008; Srirangalingam et al., 2006; Grady et al., 1997). Therefore, pay-for-performance schemes are necessary in health care because they reward physicians for higher quality, as do prices in most other markets (Robinson, 2001).

Systematic reviews of pay-for-performance schemes show that pay-for-performance contracts can influence physician behavior and improve the quality of provided primary care services (Van Herck et al., 2010; Chaix-Couturier et al., 2000). However, their effect on behavior is often complex and limited (Christianson et al., 2008) and is dependent on the program's design choices and the context (Van Herck et al., 2010). Some of the design choices that have proved to influence the effectiveness of pay-for-performance program include (1) quality goals and targets, (2) pay-for-performance incentive, and (3) program communication and implementation.

With regard to quality goals and targets, process indicators generally produced higher improvement rates in comparison with intermediate measures and outcome goals. In addition, programs that involved stakeholders in target selection and definition provided better results (Gilmore et al., 2007; Amundson et al., 2003; Chung et al., 2003; Larsen et al., 2003) in comparison with those that did not (Van Herck et al., 2010). Pay-for-performance incentives that were positive in nature (de Brants et al., 2009; Rosenthal et al., 2008; Coleman et al., 2007; Fairbrother et al., 2001; Fairbrother et al., 1999; Hillman et al., 1998; Morrow et al., 1995) generated more positive effects in comparison with incentives that were designed in a competitive format, in which some doctors won and some lost (Mullen et al., 2009; Ryan, 2009; Karve et al., 2008; Glickman et al., 2007; Lindenauer et al., 2007; Young et al., 2007; Levin-Scherz et al., 2006; Morrow et al., 1995). In addition, programs focusing on individual providers (de Brants et al., 2009; Rosenthal et al., 2008; Coleman et al., 2007; Fairbrother et al., 2001) and teams (Gilmore et al., 2007; Chung et al., 2003; Greene et al., 2004) both produced positive results. However, programs that focused on hospitals produced smaller positive effects in comparison with individual and team level programs (Pearson et al., 2008; Glickman et al., 2007; Hillman et al., 1999; Hillman et al., 1998; Morrow et al., 1995; Rosenthal et al., 2005). Communication and participant awareness can also play a significant role in the success of the pay-for-performance program. Some studies have related the ineffectiveness of their programs to lack of proper communication and awareness of the program and its elements (Hillman et al., 1999; Hillman et al., 1998). Extensive and direct communication of the program (Beaulieu & Horrigan, 2005; Amundson et al., 2003; Hillman et al., 1999) and involvement of all the

stakeholders in pay-for-performance program development (Gilmore et al., 2007; Amundson et al., 2003; Chung et al., 2003; Larsen et al., 2003) are two factors that produced positive effects.

As was explained in the literature review chapter, only a few studies have examined the effect of pay-for-performance programs on referrals (An et al., 2008; Srirangalingam et al., 2006; Grady et al., 1997). Grady et al.'s study (1997) evaluated the impact of three different approaches on mammogram referrals for patients aged 50 years and older. Physicians were randomly assigned to three treatment conditions: (1) education-only condition, (2) education plus cue enhancement strategy condition, and (3) education plus cue enhancement plus feedback and reward condition (Grady et al., 1997). Physician education included the provision of information on the factors contributing to breast cancer, including the high correlation between breast cancer and age, and the positive effect of physician encouragement on mammography use. Cue enhancement strategy consisted of using mammography chart stickers and spaces for three mammography referrals and completions. The reward was an amount of money determined by the percentage of referred patients in a specified period of time. For example, the physician would receive $50 for a 50% referral rate. 61 practices participated in the study over a three-year period resulting in a sample of 11,426 patients. While chart sticker cueing significantly increased referrals, completions, and overall compliance in comparison with education alone, the researchers did not find any improvement from performance-feedback and reward above and beyond the use of sticker cueing strategy. The authors speculate that the reward was too small and isolated to have had an effect on performance in these practices.

In another study, Srirangalingam et al. (2006) focused on the implementation of a broad-based pay-for-performance program by the UK’s National Health Service (NHS). Researchers

analyzed how referral patterns for diabetes care changed after the introduction of the new program. While the study found no significant improvement in the total number of referrals to secondary care 6 months after the implementation of the new pay-for-performance program, the thresholds for referral of diabetic patients significantly decreased during this six month period. This suggested the provision of a higher quality service indicating that physicians were acting more vigorously upon poor glycaemic control after the implementation of the new reward system (Srirangalingam et al, 2006).

Finally, a recent study done by An et al. (2008) used a randomized trial to compare the effect of a pay-for-performance program on the referrals to tobacco quitline services. Tobacco quitline services provide evidence-based methods for stopping tobacco usage. The pay-for-performance program offered $5000 if 50 patients were referred to tobacco quitline services. The study was a randomized trial that compared usual care in 25 clinics with the pay-for-performance program in 25 clinics. The results showed that clinics under the pay-for-performance program referred 11.4% of their patients. The rate was significantly higher than the 4.2% referral rate for usual care clinics (An et al., 2008).

In this study, we examine the effect of a pay-for-performance payment scheme on the medical professionals' referral rate and referral appropriateness. We define referral appropriateness based on the average of the urgency ratings of a panel of specialists for each patient case description. The closer the medical doctor’s urgency rating for each case is to the average of the ratings of the panel of specialists, the more appropriate their referral decision is to the panel’s decision. We hypothesize that a pay-for-performance policy will result in

significantly fewer referrals to specialists and more appropriate referral decisions in comparison with the fixed-payment policy.

Hypothesis 3. Medical doctors who work under the pay-for-performance policy may refer significantly fewer patients to specialists in comparison with those who work under the fixed payment policy.

Hypothesis 4. Medical doctors who work under the pay-for-performance policy may refer more appropriately in comparison with those who work under the fixed payment and fundholding policies.

6.4 Experimental Design

We designed an experiment to examine the effect of a change in GPs' method of remuneration (remuneration type: fundholding, pay-for-performance, and fixed payment) on referral rate and referral appropriateness. To test our hypotheses, we conducted an online survey with family practice residents as participants. 51 family practice residents participated in the survey and were randomly assigned to three treatment conditions based on the type of remuneration method (3 types of fundholding, pay-for-performance and fixed payment). The design of the experiment was a mixed-method matched and random assignment. The study was done online through Qualtrics software. The software randomly assigned participants to one of the three treatments. Participants were told that the purpose of the study was to understand the factors affecting priority setting in the referral process from GPs to specialists.

The study was done online through Qualtrics software. Qualtrics is a global data collection and analysis provider for research and data collection in different areas including market research, voice of customer, employee performance, and academic research. It provides a

platform for data collection and analysis (http://www.qualtrics.com/blog/about-qualtrics/). The University of Calgary is one of the universities that uses the Qualtrics platform for quantitative and qualitative research purposes.

In order to manipulate the type of payment in each treatment, we provided a description of the payment method through which the participants would be remunerated for their participation in the experiment. After reading the description about the payment method, participants were presented with 12 patient case descriptions. The case descriptions were completely anonymous. They consisted of 6 cases in each area of rheumatology and respirology. The rheumatology cases were developed by the Western Canada Waiting List group (WCWL) and provided to us by Dr. Tom Noseworthy, Professor of Health Policy and Management in the Department of Community Health Sciences and Institute for Public Health, University of Calgary. The Respirology cases were developed and provided to us by Dr. Sachin Pendharkar, Assistant Professor in the University of Calgary. For each patient case description, participants answered three questions: (1) whether they would refer the patient to the specialist, (2) the urgency or relative priority of the patient for referral on a scale from 0 (Not Urgent) to 10 (Very Urgent), and (3) a brief explanation about the reasons for their referral or non-referral decision. The 0-10 scale was a discrete version of the Visual Analogue Scale (VAS) used as the Priority-Referral Score (PRS) by the Western Canada Waiting List (WCWL) group for rating the relative urgency of referrals by general practitioners (De Coster et al., 2007). Appendix C presents all parts of the experiment including the survey description, preliminary questionnaire, case test, and after-the-test questionnaire.

6.5 Measuring Referral Appropriateness

We measured referral appropriateness using a variable called participant's overall error measure. In order to calculate the overall error measure, first we measured error per case per participant. Error for each case *i*, was measured by taking the difference between the participant's urgency rating which is on a 0-10 scale and the average of the panel of specialists' urgency ratings for the same case *i*. We call this value e_i which is the participant's error value on case *i*. Then, we calculated the overall error measure by summing the absolute values of errors for all the participant's referral decisions: $\sum ABS(e_i)$, where $ABS(e_i)$ is the absolute value of case e_i. The closer the participant's urgency ratings are to the average of the ratings of the panel of specialists, the smaller is the overall error measure in the participant's referral decision and the smaller the overall error level, the more appropriate their referral decision is.

6.6 Manipulating Payment Method

In order to manipulate payment method in each treatment condition, we provided an explanation of the payment method at the beginning of the case test section of our experiment. In the fixed pay treatment, the participants were paid $25 for their participation. Participants' payment was not dependent on their referral performance in this treatment. However, the fact that participants knew that they would receive performance feedback after the survey could psychologically motivate them to provide answers after giving careful thought to the survey questions. If they answered questions based on mere guessing, they would feel dissatisfied with themselves when they received specialists' feedback on their performance. Table 6.1 provides the remuneration method descriptions for different treatments.

In the fundholding treatment, participants received $25 as their participation fee. However, they were provided with an explanation about a budget that was allocated to their practice. The fund was $1000 for 10 patients. When they referred a patient to a specialist, the participant paid an amount of $100 out of the budget. If any of the participant's budget was left over at the end, 5% of the remainder would be given to the participant's practice as a medical resident. As an example, if the participant referred 4 patients, then $20 would be given to the participant's practice. If the participant faced a budget deficit, the participant would still receive the $25 participation fee.

In the pay-for-performance treatment, participants were paid based on their performance. The participants were told that their payment depended on their responses as compared with a panel of specialists' responses on all case descriptions. In addition to the $25 participation fee, they were able to receive an extra amount of a maximum of $25 based on how well their ratings matched the average ratings of a panel of specialists who had reviewed the same cases. If their ratings completely mismatched the panel's ratings, they would only receive the $25 participation fee. In other words, in the most extreme case, if a participant's urgency ratings on all the cases were 0 while the average urgency ratings of the panel of specialists on each case was 10, the amount of additional payment would be 25*1/(1+10*10) which is almost equal to $0.25. Therefore, the participant will receive a total $25.25.

The additional amount of a maximum of $25 in the pay-for-performance treatment was calculated using the overall error measure for each participant. Overall error measure is defined as the sum of absolute difference between the participant's urgency ratings and the specialists' average urgency ratings on each case: $\sum ABS\ (e_i)$. The additional payment was based on the

overall error measure and was formulated as: 25* 1/ (1+∑ABS (e_i)). If the participant's ratings completely matched the panel of specialists' average ratings, the participant would receive an additional amount of $25 for his/her performance. If the participant's ratings completely mismatched the panel's average ratings, the participant would not receive any additional amount for performance.

Table 6.1 Payment scheme manipulation

Remuneration Method Descriptions
a) Fixed Payment Treatment The following case describes one of your patients who you may consider referring to a specialist. Please read the case and answer the questions below.
b) Fundholding Treatment In this section you will be presented with a set of cases that describe your patients. You may consider referring some of them to a specialist. Assume that you have a budget of $1000 for the referral of these patients. Any time you refer a patient to a specialist, you will pay $100 out of the budget. If any of your budget is left over at the end, 5% of the remainder will be given to your practice as a medical resident. Please read the cases and answer the related questions.
c) Pay-for-Performance Treatment In this section you will be presented with a set of cases that describe your patients. You may consider referring some of them to a specialist. In addition to your participation fee, you can receive an extra amount of a maximum of $25 based on how well your ratings match the average ratings of a panel of specialists who have reviewed the same cases. Please read each case carefully and answer the related questions.

6.7 Variables

6.7.1 Dependent Variables

Table 6.2 depicts all the variables that were measured and included in this study. The study has three main dependent variables: the total number of referrals from the participant as the GP to specialist, the average urgency rating for each participant, and the referral appropriateness rating. The survey instrument had 12 patient case descriptions. For each patient case description,

the participants answered whether they would refer the patient to a specialist (yes/no question) and how urgent it was for the patient to visit the specialist on a 0-10 scale. The total number of referrals was calculated by counting the total number of positive responses by the participant on the first question. The average urgency rating is calculated by finding the average of urgency ratings on all the patient case descriptions. In order to examine the degree of referral appropriateness, we used the overall error measure described above, which is the difference between the participant's urgency ratings and the specialists' average urgency ratings on each case. The lower the overall error measure, the more appropriate the participant's referral decisions are.

6.7.2 Independent Variables

The study has one main independent variable: remuneration method (Pay) that has three levels of fixed payment, fundholding, and pay-for-performance. We used dummy variable coding (0/1) with two categories for fundholding and pay-for-performance payment methods. Fixed payment was used as the reference category.

Table 6.2 Variables

Variable Name	Variable	Measure	Variable Type
Payment type (Payment)	Independent	Type of payment scheme	Nominal with three levels : 1 = Fixed pay; 2 = Pay-for-Performance; 3 = Fundholding
Fundholding	Independent	Fundholding payment type	Nominal with two levels: 0.5 = Fundholding; -0.5 = Other
Pay-for-Performance	Independent	Pay-for-Performance payment type	Nominal with two levels: 0.5 = Pay-for-Performance; -0.5 = Other
Confidence	Independent	Confidence in referral decision making	Scale (0-7)
Referral rate (RefRate)	Dependent	Number of referrals	Scale (0-12)
Urgency (AvgUrgency)	Dependent	Average of urgency ratings on the 12 patient case descriptions	Scale (1-10)
Mood	Dependent	Two items selected from PANAS (Positive and Negative Affect Scale) mood scale	Scale (1-7)
Knowledge (Know)	Covariate	Level of knowledge in the areas of rheumatology or respirology depending on the test type	Scale (1-7)
Age	Covariate	Participant's age	Ordinal based on age group
Gender	Covariate	Participant's gender	Nominal with two levels: 1 = Female; 0 = Male
Year of program (Year)	Covariate	Participant's year of residency program	Scale
Location of program (Rural)	Covariate	Participant's location of residency program	Nominal with two levels: 1 = Rural; 0 = Urban
Primary Language (English First)	Covariate	Participant's primary language	Nominal with two levels: 1 = English; 0 = Other
Risk	Covariate	Selected items from Holtgrave et al.'s (1990) multidimensional risk scale	Scale (1-7)
Loss Aversion (LossAverse)	Covariate	One item used in Nightingale's studies (1987a,b)	Nominal with two levels: 0 = Risk seeking; 1 = Loss Averse

6.7.3 Control Variables

In addition to the primary measures of interest, we collected data on a set of other variables in order to control for variance due to the effect of these variables on our dependent variable based on the literature on factors affecting referral rates and referral appropriateness.

Based on the referral literature, GP characteristics can have a significant effect on referral decision making. As was mentioned in the previous chapters, a GP's knowledge, confidence, risk seeking preferences, and loss aversion can influence their referral decision making. For the readers' convenience, we explain these effects here again. First, the literature shows that GP knowledge and confidence towards a specific area of specialty may either decrease (Newton et al., 1991) or increase (Reynolds et al., 1991; Morrell et al., 1971) referrals in a specific area of specialty. In order to measure knowledge level in each area of specialty, we asked participants to indicate their level of knowledge in both rheumatology and respirology prior to the case test on a scale from 0 (not at all knowledgeable) to 7 (very knowledgeable). In order to measure confidence, we asked participants to rate their confidence level using one item: “Please rate your confidence in your ability to determine whether a patient of yours should be referred to a specialist in the two areas of rheumatology and respirology”; 0 = not at all confident; 7 = very confident.

Second, studies have found that GP risk seeking preferences and degree of loss aversion can significantly impact different clinical factors such as laboratory usage (Holtgrave et al., 1990) and referral to hospital (Nightingale et al., 1988; Nightingale, 1987). In order to measure participants' risk attitude, we used a multidimensional scale because several studies have shown that risk attitude is multidimensional (Jackson et al., 1978; Weber, 1988). Similarly, risk attitude in medical decision making must consider different aspects of the individual’s risk seeking preferences (Holtgrave et al., 1990). Therefore, using a multidimensional scale, we measured risk in a number of different domains including medical, financial, social, and physical domains.

Table 6.3 and 6.4 provide the scale and binary items of the risk attitude measure used in this study.

The risk attitude measurement instrument consisted of an 8-item measure, which participants voluntarily completed after they finished the case test. The survey instrument was placed after the case test in order to avoid any confounding effect of individual's performance in the case test with risk measurement. The risk attitude instrument was a reduced version of Holtgrave et al.'s (1990) multidimensional risk scale. It comprised items both from Nightingale's risk measurement in medical domain (Nightingale et al., 1988; Nightingale, 1987a, 1987b) and Jackson's standardized risk measurement method (Jackson et al., 1978) that covered risk attitudes in financial, physical, and social domains. In order to keep the survey length short, we decided to use an abbreviated version of the risk attitude instrument used by Holtgrave et al. in this study.

Table 6.3 Likert scale items of risk measure

Seven-point Likert scale items of risk measure 1(Not at all true)-7 (Very true)
I am outspoken. (Feelings)
I am concerned about getting hurt. (Danger)
I enjoy an element of physical danger. (Danger)
I am careful about not hurting people's feelings. (Feelings)
I am generally risk-averse when dealing with patients (RiskPatient).

Table 6.4 Binary choice items of risk measure

Binary choice items of risk attitude measure
I would not speak out in public on unpopular issues unless I were fairly certain of my position. (Public) True/False
If asked to choose between the two following therapies for your patient, which one would you select? (Gain) ☐ Therapy A, which **definitely** gives the patient five more years of life than the average longevity ☐ Therapy B, which gives a 50:50 chance of zero or ten years of additional life for your patient
If asked to choose between the two following therapies for your patient, which one would you select?(Loss) ☐ Therapy A, which the patient definitely lives five years less than the average person ☐ Therapy B, which gives a 50:50 chance of losing zero or ten years of additional life for the patient

Third, we controlled for the geographic location of the participants' residency program by asking participants the location of their residency program (urban versus rural). Based on the literature, there is regional variation in the management of medical problems (Pilote et al., 1995; Notzon et al., 1987) and in the use of medical and surgical services such as services that are made available through the referral process (Birkmeyer et al., 1998; Wilkin & Smith, 1987; Chassin et al., 1986; Wennberg, 1986; Boyle, 1985; McPherson et al., 1985). GP's geographic location is an important indicator of referral rates (Langley et al., 1997; 1990; Jones, 1987); rural GPs have significantly lower referral rates in comparison with urban GPs (Langley et al., 1997; Madeley et al., 1990). Finally, we collected participants' demographic data including age group, gender, ethnicity, and whether English is their first language.

6.7.4 Other Dependent Variables

In addition to risk attitude, we measured participants' mood using a 2-item scale selected from the PANAS scale right after the completion of the case test section (PANAS; Watson et al., 1988). On a 7-point scale, we asked participants to indicate how much they agreed with each of the following two sentences: 1. I felt **cheerful** while I was completing the test and 2. I felt **upset** while I was completing the test (1 = not at all true; 7 = very true). We measured mood to test the likelihood of the impact of our manipulation for payment method on participant's mood and to eliminate other explanations for the observed referral decisions and urgency ratings.

The results of a Univariate Analysis of Variance on the effect of payment type on mood showed that our manipulation for the payment did not have a significant effect on mood (F = 1.406; p = 0.582). Therefore, the changes in referral rates and urgency ratings could not be attributed to changes in the participants' mood.

6.8 Participants and Procedures

The online survey was emailed to family practice residents through the Departments of Family Medicine in seven Canadian universities including University of Calgary, University of British Columbia, University of McGill, University of Ottawa, University of Manitoba, University of Saskatchewan, and Memorial University of Newfoundland. The approximate sample of residents who received the email was 900 people among which 51 family practice residents participated in this study. They were randomly assigned to the conditions of a one-way between subjects factorial design. The survey was sent to residents four times in October, and December, 2012, January and February, 2013. The highest response rate was obtained in the third round in January 11, 2013. This can be due to the fact that the survey was sent right after

the Christmas holidays when participants were back from holiday breaks, so they were fresh and less busy in comparison with other rounds.

At first, participants read a survey description which consisted of information on the survey purpose, its content, length of time it took to complete the survey, and the payment level. The purpose of the survey as explained to participants was “to understand the factors affecting priority setting in the patient referrals from general practitioners to specialists”. Participants in the fixed payment and fundholding treatments were told that they would receive $25 for their participation in the survey. Participants in the pay-for-performance treatment were told that they would receive a maximum of $25 in addition to their $25 participation fee, the amount of the extra payment dependent on their performance. Appendix C.1 provides the complete survey description for each of the three treatment conditions.

In section 1, participants answered a preliminary questionnaire including questions regarding their level of knowledge and confidence in answering questions in two areas of rheumatology and respirology. Appendix C.2 presents the preliminary questionnaire used in this study.

Section 2 included the case test which consisted of 12 patient case descriptions in the two areas of rheumatology and respirology. In order to explain the remuneration system, we provided a written description of the payment method in each condition which was presented in table 2 of the Manipulating Payment Method section. After reading the payment description, participants were guided to continue with reading the patient case descriptions and answering the related questions. Appendix C.3 presents a shortened version of the case test section for the three treatment condition.

Finally, in section 3, participants completed an after-the-test questionnaire that consisted of questions on mood (PANAS, Watson et al., 1988), risk attitude, geographic location of residency program, and demographic data. Appendix C.4 presents the after-the-test questionnaire.

6.9 Main Statistical Analyses and Results

Descriptive statistics of the data showed that 47.6% of the participants were female and 52.4% were male. In addition, almost half of the participants (54%) were working in the urban areas. There was not a significant difference in the number of rural versus urban participants in the four treatment conditions. Participants were all either in the first or second year of their residency program. There was not a significant difference as regards with year, age and English language among the three treatment conditions.

In order to be able to test the difference in referral behavior between high-confidence and low-confidence individuals in different treatments, we converted the continuous variable "Confidence" to a categorical variable called "RefConf". Categorizing the confidence variable allowed us to test the effect of relationship type on referral rate for high-confidence and low-confidence participants separately. The RefConf value for those whose confidence level was less thsan 5 was coded as -0.5, while the RefConf value for those whose confidence level was more than or equal to 5 was coded as 0.5. Table 6.5 provides the descriptive statistics for all the measured variables in total and in each treatment condition.

Table 6.5 Descriptive statistics

Descriptive Statistics	All Treatments			Fixed-pay			Pay-for-performance			Fundholding Treatment		
Variable	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**	**Mean**	**Std Dev.**	**N**
RefRate	5.20	1.22	51	5.29	1.31	17	5.24	1.03	17	5.06	1.34	17
AvgUrgency	4.05	1.26	51	3.84	1.30	17	4.13	1.29	17	4.18	1.23	17
RefError	29.37	6.32	51	29.80	4.99	17	28.80	6.02	17	29.51	7.95	17
Mood	4.36	1.18	51	4.38	1.34	17	4.38	1.17	17	4.32	1.07	17
Payment	2.00	0.82	51	1.00	0.00	17	2.00	0.00	17	3.00	0.00	17
Fundholding	-0.17	0.48	51	-0.50	0.00	17	-0.50	0.00	17	0.50	0.00	17
Pay-for-Performance	-0.17	0.48	51	-0.50	0.00	17	0.50	0.00	17	-0.50	0.00	17
Confidence	4.66	0.98	50	4.69	0.89	16				4.62	0.94	17
Know	3.63	0.75	51	3.62	0.74	17	3.53	0.60	17	3.74	0.90	17
RefConf	0.00	0.51	50	0.00	0.52	16	-0.03	0.51	17	0.03	0.51	17
Risk	2.44	0.95	51	2.27	0.64	17	2.89	1.35	17	2.17	0.51	17
LossAverse	0.47	0.50	49	0.56	0.51	16	0.29	0.47	17	0.56	0.51	16
Rural	0.29	0.46	51	0.24	0.44	17	0.35	0.49	17	0.29	0.47	17
Gender	0.67	0.48	51	0.71	0.47	17	0.65	0.49	17	0.65	0.49	17
Age	1.51	0.51	47	1.44	0.51	16	1.60	0.51	15	1.50	0.52	16
Year	1.35	0.56	51	1.41	0.62	17	1.35	0.61	17	1.29	0.47	17
EnglishFirst	0.90	0.30	51	0.94	0.24	17	0.88	0.33	17	0.88	0.33	17
Valid N			44			14			15			15

To test the first hypothesis regarding the effect of a fundholding payment scheme on referral rate, we ran multiple regression models with referral rate as our dependent variable. In order to test the interactions between independent variables of interest, we included the two-way interaction between the Fundholding and RefConf to our multiple regression models.

In order to avoid multicollinearity, we added the variables to the model in a step-by-step format. Table 6.6 shows the results of the multiple regression analysis in the step-by-step format. In the first step, we included all the control variables (Model 1). In the second step, we included the main independent variable of Fundholding and RefConf (Model 2). We did not find any significant effect of Fundholding on Referral Rate to support our first hypothesis. However, the results were in the direction that we had expected. Figure 6.1 shows the direction of the relationship between fundholding and referral rate. In the third step, we added the two-way interaction between Fundholding and RefConf (Model 3). Though we did not find a significant main effect for Fundholding, we found a significant interaction between Fundholding and RefConf (β = -2.299; p = 0.032).

The fact that the overall model F statistic is not significant (F = 0.983; p = 0.485) makes it difficult to interpret the results of the regression analysis. The insignificance of the F statistic can be related to our small sample size.

In order to increase the power of the test and to get a significant overall model F statistic, we ran Model 4 with no control variables included in the model. The interaction was significant in this parsimonious model as well (β = -1.958; p = 0.038). But, we did not get a significant F statistic for the overall regression model even in this parsimonious model (F = 1.705, p = 0.188). Models 5 to 8 show the follow-up tests for high-confidence and low-confidence groups. The

follow-up tests for the interaction show that those residents who have high referral confidence refer significantly fewer patients under the fundholding treatment rather than the fixed payment treatment.

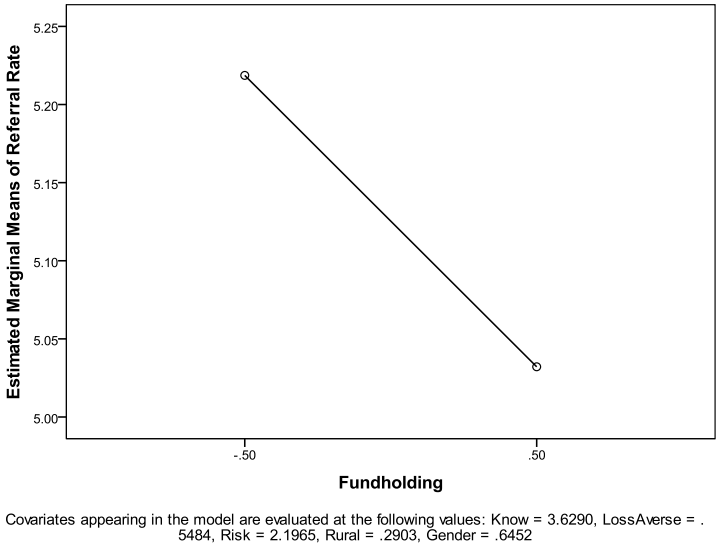


Figure 6.1 Fundholding and referral rate

Table 6.6 Multiple regression_dependent variable: referral rate

	Dependent Variable: Referral Rate							
	Main Test				Follow-up Tests			
					High RefConf		Low RefConf	
	Model 1	Model 2	Model 3	Model 4	Model 4	Model 5	Model 6	Model 7
Constant	4.139 (0.016)	4.186 (0.027)	3.505 (0.045)	5.198 (0.000)	5.211 (0.153)		2.300 (0.158)	
Knowledge (1-7)	0.254 0.158 (0.420)	0.285 0.178 (0.440)	0.400 0.249 (0.248)		0.075 0.041 (0.907)		1.242 0.750 (0.013)	
Loss Aversion (0/1)	-0.465 -0.180 (0.415)	-0.434 -0.166 (0.483)	-0.566 -0.216 (0.326)		-0.188 -0.072 (0.824)		-0.951 -0.368 (0.177)	
Risk	0.073 0.030 (0.880)	0.005 0.002 (0.992)	0.055 0.023 (0.911)		-0.418 -0.124 (0.696)		-0.155 -0.081 (0.734)	
Rural (0/1)	-0.441 -0.155 (0.436)	-0.428 -0.149 (0.494)	-0.393 -0.137 (0.495)		0.465 0.136 (0.650)		-0.834 -0.322 (0.185)	
Gender (0/1)	0.395 0.146 (0.512)	0.366 0.134 (0.581)	0.735 0.270 (0.250)		0.975 0.340 (0.370)		-0.275 -0.106 (0.711)	
Fundholding		-0.313 -0.120 (0.567)	-0.186 -0.071 (0.712)	-0.229 -0.087 (0.615)	- 1.503* -0.574 (0.086)	- 2.417* -0.461 (0.062)	0.695 0.272 (0.276)	0.750 0.285 (0.285)
Referral Confidence (RefConf)		-0.056 -0.021 (0.929)	-0.146 -0.056 (0.799)	0.146 0.056 (0.748)				
Fundholding*RefConf			- 2.299** -0.441 (0.032)	- 1.958** -0.373 (0.038)				
N	31	30	30	32				
R Squared (%)	7.3	8.7	26.3	15	33.5	21.3	69.1	8.1
Adjusted R Squared (%)	-10.5	-19	-0.5	6.2	-6.4	16	42.6	1.5
F Statistics	0.410	0.314	0.983	1.705	0.839	4.055*	2.610	1.235
P value	0.837	0.940	0.475	0.188	0.567	0.062	0.118	0.285

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

Figure 6.2 shows the interaction between payment type (fundholding versus fixed pay) and confidence as a categorical variable and their interaction effect on referral rate. As the figure shows, participants who had high confidence referred more patients to participants under the fixed-pay method in comparison with the fundholding method. On the contrary, participants who had low confidence referred more patients under the fundholding method in comparison with the fixed-pay method.

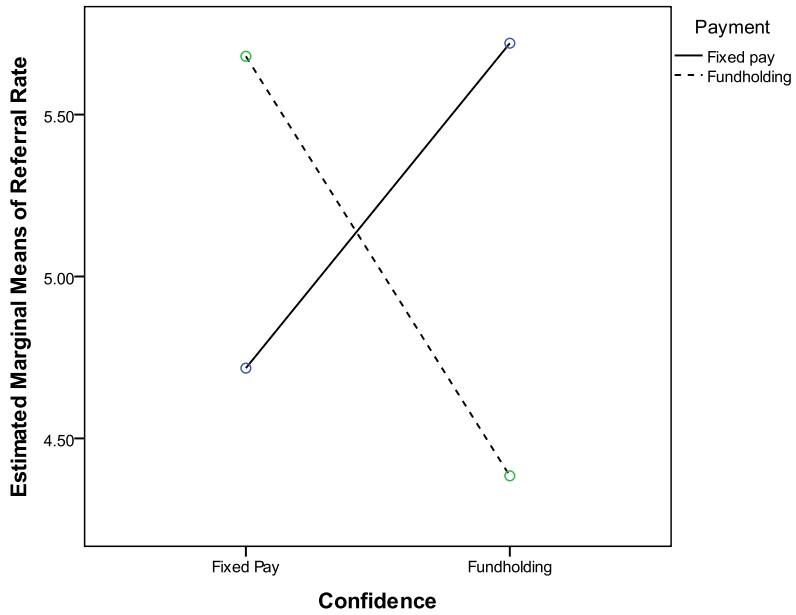


Figure 6.2 Fundholding x confidence interaction and referral rate

From a statistical point of view, the insignificance of the overall F statistic means that the linear regression model is not a good fit for the data. One reason can be the existence of a nonlinear relationship between the independent and dependent variables (Neter & Wasserman, 1974). A simple regression plot of the continuous variable confidence and referral rate showed a

significant quadratic relationship between the two variables (F = 4.436; p = 0.017). Figure 6.3 shows the quadratic relationship between Confidence and Referral Rate. Thus, we added the quadratic form of Confidence, ConfQuad, to the model. Table 6.7 provides the regression models with Confidence and ConfQuad variables substituting the categorical variable, RefConf in the new regression models. Models 3 and 4 indicate that the variables Fundholding, Confidence, ConfQuad, and the interaction between Confidence and Fundholding are all significant when the quadratic term is added to the model. These results confirm the existence of a significant quadratic relationship between confidence and referral rate. In other words, those who have low or high confidence refer significantly fewer patients in comparison with those who have medium confidence level.

This result may be just an artifact of our small sample size or due to confidence measurement in this study: Each participant reported their confidence for both rheumatology and respirology areas separately. Confidence was then calculated by taking the average confidence in both areas of respirology and rheumatology.

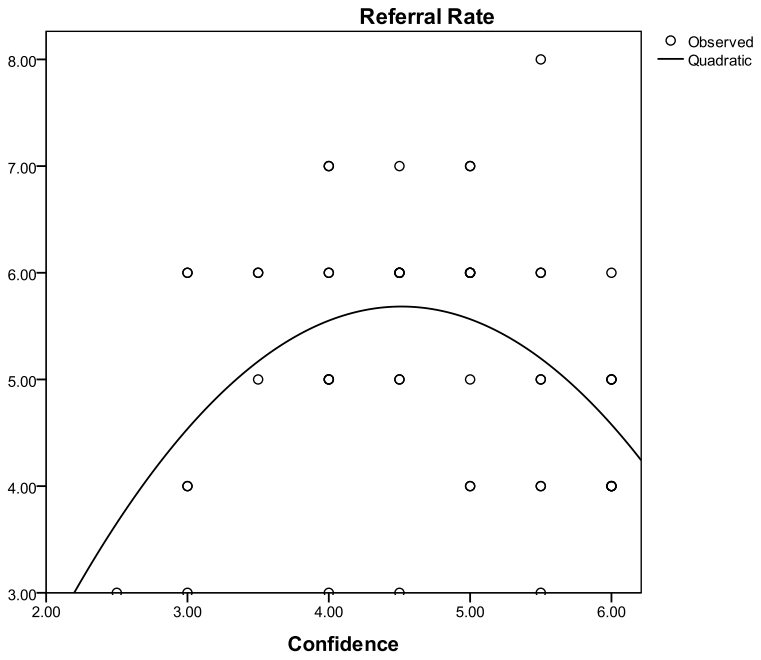


Figure 6.3 Quadratic relationship between confidence and referral rate

Table 6.7 Multiple regression_dependent variable: referral rate

	Dependent Variable: Referral Rate							
	Main Test				Follow-up Tests			
					Fixed Rate		Fundholding	
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	4.139 (0.016)	-10.577 (0.020)	-11.061 (0.0112)	-11.254 (0.004)	-8.729 (0.322)	-9.137 (0.190)	-8.528 (0.148)	
Knowledge (1-7)	0.254 0.158 (0.420)	-0.148 -0.092 (0.660)	-0.085 -0.053 (0.791)		-0.276 -0.150 (0.712)		0.245 0.173 (0.573)	
Loss Aversion (0/1)	-0.465 -0.180 (0.415)	-0.389 -0.148 (0.440)	-0.290 -0.111 (0.547)		-0.302 -0.112 (0.712)		-0.093 -0.037 (0.894)	
Risk	0.073 0.030 (0.880)	-0.435 -0.179 (0.333)	-0.367 -0.151 (0.393)		-0.943 -0.439 (0.171)		0.994 0.331 (0.266)	
Rural (0/1)	-0.441 -0.155 (0.436)	-0.578 -0.201 (0.273)	-0.527 -0.183 (0.294)		-0.421 -0.139 (0.686)		0.011 0.004 (0.987)	
Gender (0/1)	0.395 0.146 (0.512)	0.248 0.091 (0.649)	0.322 0.118 (0.537)		-0.290 -0.102 (0.774)		0.483 0.187 (0.509)	
Fundholding		0.088 0.034 (0.843)	3.586* 1.374 (0.083)	4.150** 1.579 (0.028)				
Confidence		8.105*** 6.353 (0.001)	8.008*** 6.277 (0.001)	7.457*** 5.671 (0.000)	7.507 4.990 (0.121)	5.946* 3.922 (0.068)	5.358 4.856 (0.133)	8.310*** 7.074 (0.000)
ConfQuad		- 0.895*** -6.315 (0.001)	- 0.876*** -6.179 (0.001)	- 0.807*** -5.538 (0.000)	-0.772 -4.637 (0.135)	- 0.591* -3.527 (0.096)	-0.639 -5.201 (0.103)	- 0.949*** -4.802 (0.000)
Fundholding* Confidence			-0.744* -1.374 (0.084)	-0.857** -1.561 (0.030)				
N	31	30	30	32	14	15	15	16
R Squared (%)	7.3	45.1	52.6	47.7	58.9	33.8	69.1	63.3
Adjusted R Squared (%)	-10.5	25.2	32.3	40.2	17.8	23.6	42	58.1
F Statistics	0.410	2.263*	2.588**	6.381***	1.432	3.322*	2.553	12.8***
P value	0.837	0.062	0.035	0.001	0.324	0.068	0.106	0.001

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

In order to test the second hypothesis on the effect of fundholding on referral urgency, we ran another multiple regressions in a step-by-step format as we did for referral rate. The results of the analysis did not show a significant effect of fundholding on average referral urgency (β = 0.326, p = 0.47). Table 6.8 provides the results of the multiple regression analysis on the effect of fundholding on referral urgency. In this table, we have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model.

Table 6.8 Multiple regression_dependnet variable: referral urgency

Dependent Variable: Referral Urgency				
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	2.255 (0.132)	2.064 (0.196)	1.949 (0.238)	
Knowledge (1-7)	0.527 0.340 (0.070)	0.476 0.310 (0.147)	0.496 0.323 (0.144)	
Loss Aversion (0/1)	0.005 0.002 (0.992)	-0.007 -0.003 (0.989)	-0.029 -0.012 (0.958)	
Risk	-0.127 -0.054 (0.772)	0.037 0.016 (0.937)	0.045 0.020 (0.924)	
Rural (0/1)	-0.537 -0.195 (0.294)	-0.634 -0.231 (0.252)	-0.628 -0.229 (0.265)	
Gender (0/1)	0.350 0.134 (0.518)	0.469 0.180 (0.421)	0.532 0.204 (0.388)	
Fundholding		0.423 0.170 (0.380)	0.445 0.178 (0.369)	0.269 0.108 (0.556)
Referral Confidence (RefConf)		0.032 0.013 (0.953)	0.017 0.007 (0.976)	0.447 0.180 (0.330)
Fundholding*RefConf			-0.388 -0.078 (0.694)	0.222 0.045 (0.807)
N	31	30	30	32
R Squared (%)	19.5	23.5	24	4.8
Adjusted R Squared (%)	4	0.2	-3.6	-0.5
F Statistics	1.259	1.008	0.870	0.486
P value	0.311	0.451	0.556	0.694

To test the third hypothesis regarding the effect of pay-for-performance scheme on referral rate, we ran a Multiple Regression Analysis with referral rate as our dependent variable. Similar to the fundholding method, it was predicted that the interaction between the participant's referral confidence and pay-for-performance may affect referral behavior among participants. In order to test the interactions between confidence and payment type, we included the two-way interaction between pay-for-performance and confidence to our multiple regression models.

Similar to the previuos regression models, we entered the variables to the model in a step-by-step format. The results of the step-by-step multiple regression analysis revealed a marginally significant interaction between pay-for-performance and RefConf (β = -1.799; p = 0.051). However, since the F statistic for the overall regression model is not significant, it is difficult to interpret the significant interaction term in the main regression model. Table 6.9 shows the results of the multiple regression models for the effect of pay-for-performance on referral rate. Figure 6.4 depicts the interaction effect of pay-for-performance and confidence on referral rate. Similar to the fundholding method, participants who had high confidence referred fewer patients under the pay-for-performance method in comparison with the fixed-pay method. On the contrary, participants who had low confidence referred fewer patients under the fixed-pay method in comparison with the pay-for-performance method.

Table 6.9 Multiple regression_dependent variable: referral rate

Dependent Variable: Referral Rate								
	Main Test				Follow-up Tests			
					High RefConf		Low RefConf	
Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 4**	**Model 5**	**Model 6**	**Model 7**
Constant	4.451 (0.008)	4.836 (0.011)	4.568 (0.012)	5.267 (0.000)	5.102 (0.066)	5.438 (0.000)	4.203 (0.291)	5.097 (0.000)
Knowledge (1-7)	0.250 0.140 (0.469)	0.146 0.082 (0.720)	0.182 0.101 (0.636)		0.295 0.173 (0.571)		0.213 0.089 (0.809)	
Loss Aversion (0/1)	-0.126 -0.054 (0.779)	-0.106 -0.044 (0.839)	-0.020 -0.008 (0.967)		-0.092 -0.041 (0.890)		0.278 0.113 (0.770)	
Risk	-0.121 -0.112 (0.574)	-0.133 -0.123 (0.584)	-0.132 -0.123 (0.562)		-0.411 -0.310 (0.366)		-0.095 -0.102 (0.790)	
Rural (0/1)	-0.178 -0.071 (0.707)	-0.154 -0.061 (0.767)	0.012 0.005 (0.980)		-0.421 -0.163 (0.607)		0.169 0.068 (0.839)	
Gender (0/1)	0.491 0.200 (0.311)	0.504 0.204 (0.327)	0.543 0.220 (0.264)		0.634 0.275 (0.389)		0.273 0.106 (0.759)	
Pay-for-Performance (P4P)		0.068 0.029 (0.894)	0.083 0.035 (0.863)	-0.090 -0.039 (0.824)	-0.565 -0.253 (0.453)	-0.875 -0.395 (0.133)	1.069 0.443 (0.231)	0.694 0.293 (0.253)
Referral Confidence (RefConf)		0.295 0.126 (0.577)	0.349 0.149 (0.484)	0.340 0.146 (0.404)				
P4P*RefConf			-1.799* -0.383 (0.051)	-1.569* -0.338 (0.061)				
N	31	31	31		15	15	15	16
R Squared (%)	10.3	10.4	24.3	13.4	35.3	15.4	19.1	8.6
Adjusted R Squared (%)	-11.2	-15.7	-2	4.4	-7.8	9.3	-34.8	2.5
F Statistics	0.480	0.398	0.924	1.490	0.818	2.541	0.355	1.410
P value	0.817	0.894	0.516	0.238	0.582	0.133	0.890	0.253

Note: We have provided the unstandardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

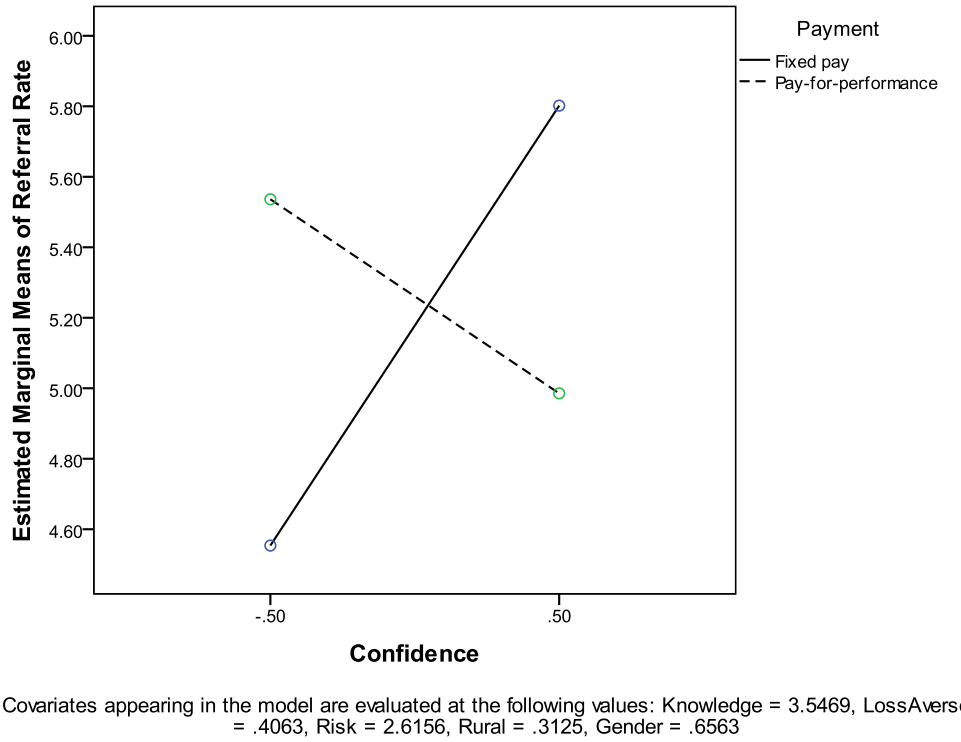


Figure 6.4 Pay-for-Performance x confidence interaction and referral rate

Similar to the fundholding treatment, we substituted RefConf with Confidence and ConfQuad variables in order to solve the issue of insignificant overall model *F* statistic. For the pay-for-performance treatment, the use of confidence and ConfQuad did not improve the overall model significance. Table 6.10 provides the results of these regression models.

Table 6.10 Multiple regression_dependent variable: referral rate

Dependent Variable: Referral Rate								
	Main Test				Follow-up Tests			
					Fixed Pay		P4P	
Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 4**	**Model 5**	**Model 6**	**Model 7**
Constant	4.451 (0.008)	3.551 (0.568)	5.059 (0.394)	1.121 (0.813)	-8.729 (0.322)	-9.137 (0.19)	14.555 (0.145)	11.825 (0.069)
Knowledge (1-7)	0.250 0.140 (0.469)	0.210 0.117 (0.646)	0.264 0.147 (0.542)		-0.276 -0.150 (0.712)		0.433 0.251 (0.477)	
Loss Aversion (0/1)	-0.126 -0.054 (0.779)	-0.094 -0.039 (0.865)	0.125 0.052 (0.816)		-0.302 -0.112 (0.712)		0.252 0.115 (0.774)	
Risk	-0.121 -0.122 (0.574)	-0.120 -0.112 (0.650)	-0.273 -0.253 (0.302)		-0.943 -0.439 (0.171)		-0.180 -0.236 (0.604)	
Rural (0/1)	-0.178 -0.071 (0.707)	-0.182 -0.072 (0.745)	-0.101 -0.040 (0.848)		-0.421 -0.139 (0.686)		0.240 0.114 (0.770)	
Gender (0/1)	0.491 0.200 (0.311)	0.481 0.195 (0.372)	0.524 -0.212 (0.304)		-0.290 -0.102 (0.774)		0.887 0.423 (0.176)	
Pay-for-Performance (P4P)		0.054 0.023 (0.920)	5.016* 2.135 (0.063)	3.683 1.584 (0.103)		5.946* (0.068)		-2.853 (0.314)
Confidence		0.440 0.340 (0.878)	-0.232 -0.179 (0.932)	1.650 1.266 (0.444)	7.507 4.990 (0.121)	-0.591* (0.096)	-4.533 -4.145 (0.257)	0.297 (0.341)
ConfQuad		-0.045 -0.312 (0.885)	0.029 0.205 (0.920)	-0.158 -1.096 (0.507)	-0.772 -4.637 (0.135)		0.446 3.694 (0.292)	
P4P*Confidence			-1.036* -2.093 (0.062)	-0.802* -1.635 (0.093)				
N	31	31	31	32	14	15		
R Squared (%)	10.3	9.3	23	13.2	58.9	33.8	33.9	8.8
Adjusted R Squared (%)	-11.2	-22.2	-8.5	0.8	17.8	23.6	-17.5	-4.2
F Statistics	0.480	0.295	0.730	1.067	1.432	3.322*	0.660	0.679
P value	0.817	0.960	0.678	0.391	0.324	0.068	0.701	0.523

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

In order to test the fourth hypothesis on the effect of payment type on referral error, we ran multiple regression models in a step-by-step format as we did for referral rate. The results of the analysis did not show a significant relationship between payment and referral error (β = -0.327, p = 0.515). Table 6.11 provides the results of the multiple regression analysis on the effect of different independent variables on referral error. However, the results were in the direction that we had predicted. Participants in the pay-for-performance treatment had lower referral error in comparison with those in the fixed pay and fundholding treatment conditions. Figure 6.5 illustrates the effect of different payment schemes on referral error.

Table 6.11 Multiple regression_dependent variable: referral error

Dependent Variable: Referral Error			
Variables	**Model 1**	**Model 2**	**Model 4**
Constant	19.065 (0.005)	33.909 (0.000)	
Knowledge (1-7)	0.387 0.047 (0.759)	0.048 0.006 (0.974)	
Loss Aversion (0/1)	1.307 0.107 (0.494)	1.315 0.106 (0.540)	
Risk	-1.395 -0.216 (0.167)	-1.371 -0.212 (0.239)	
Rural (0/1)	-0.594 0.045 (0.767)	0.704 0.053 (0.741)	
Gender (0/1)	-1.686 -0.132 (0.406)	-1.746 -0.136 (0.420)	
Pay-for-Performance		-0.452 -0.035 (0.859)	-1.056 -0.079 (0.643)
Fundholding		-0.573 -0.044 (0.812)	-0.461 -0.035 (0.839)
Referral Confidence (RefConf)		1.164 0.095 (0.594)	1.896 0.150 (0.308)
N	48	47	49
R Squared (%)	32.2	7	2.8
Adjusted R Squared (%)	19.6	-.12	-3.6
F Statistics	2.565	0.370	0.515
P value	0.05	0.930	0.674

Note: We have provided the unstandardized and standardized regression coefficients for each variable. Values in parentheses are the significance level of each variable included in the model. * significant at 0.1, ** significant at 0.05, *** significant at 0.01

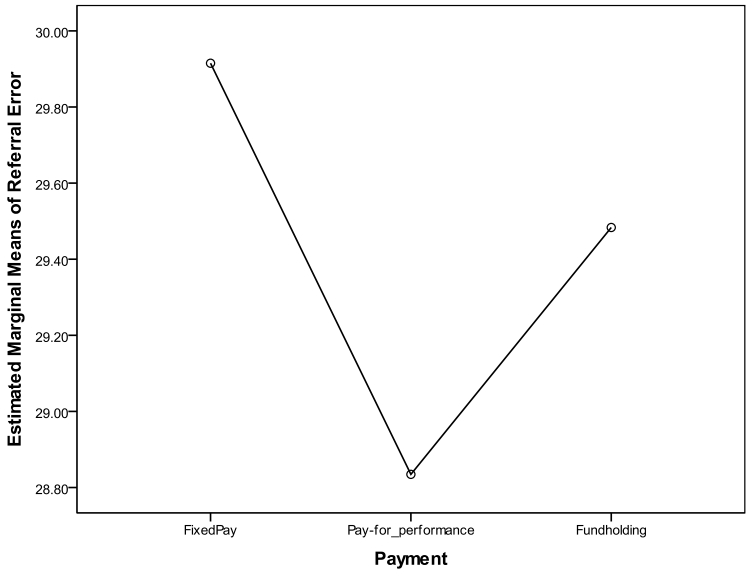


Figure 6.5 Payment schemes and referral error

6.10 Conclusion

In this study, we examined the effect of different payment schemes on referral rate and referral appropriateness. Although the overall model *F* statistics were not statistically significant, the results of the multiple regression models provided evidence on the effect of fundholding scheme on referral rate among individuals who had high level of perceived confidence. In the fundholding treatment, payment decreased referral rate among those who had had high confidence or perceived themselves capable of making informed referral decisions. However, the study did not find support for the hypothesis among those who had low levels of perceived confidence.

Confidence as addressed in this experiment referred to the participants' perception of their ability to make informed referral decisions. Since our participants were family practice residents

and did not have enough knowledge and experience with the referral decision making, it can be argued that the referral rates of residents who had higher levels of perceived knowledge and confidence can be closer to primary care physicians who gain knowledge and experience with the referral process over the years of practice. Thus, the study can give support to the hypotheses regarding the negative effect of fundholding on referral rate among primary care physicians. However, we may repeat this study with primary care physicians in a follow-up study in order to get a better insight into the effect of different payment schemes on GPs' referral behavior.

Our study has some limitations: First, the sample size is small. We predict that we may get significant results if a larger sample of family practice residents participates in our study. Second, our participants are family practice residents who have not obtained enough knowledge and experience with the referral process. In the next phase, we may repeat our survey with family physicians as participants. This may give us a better understanding of the effect of different payment schemes on referral patterns among general practitioners.

Chapter Seven: Conclusion

In this dissertation, we conducted three experiments to examine the effect of different motivational factors including monetary rewards and social relationships on performance. In the first study, on the effect of external reward and intrinsic motivation on overall task motivation and performance, we designed and conducted an experiment to test the interaction effect of performance-contingent rewards and intrinsic motivation on overall task motivation and performance. The results showed that performance-contingent rewards improved both task motivation and performance. Similar to external reward, intrinsic motivation also improved both task motivation and performance. We did not find a significant interaction between external reward and intrinsic motivation. But, the interaction was in the direction that we expected, providing evidence for the hypothesis that in the presence of both external reward and intrinsic motivation, performance may be higher than in the lack of either or both of them. This result is in contrast with cognitive evaluation theory's predictions that external reward would undermine intrinsic motivation and performance. Our results provide support for the widespread use of monetary rewards for performance improvement in different organizational settings.

In future studies, we would like to test how contextual factors such as autonomy versus control and monetary reward salience moderate the effect of external reward on motivation and performance. Based on cognitive evaluation theory, performance-contingent reward−the type of reward that is dependent on individual's performance−has the highest tendency to undermine intrinsic motivation (Deci et al., 1999). Performance-continent rewards tend to undermine intrinsic motivation due to their highly controlling nature (Deci et al., 1999). They induce people to do things that they would not do in the absence of such rewards (Deci and Ryan, 1987).

However, in a few studies where the controlling aspects of the rewards were minimized, rewards proved to maintain or even enhance intrinsic motivation (Harackiewicz et al., 1984; Ryan et al., 1983).

Contrary to CET's predictions, we hypothesize that performance-contingent rewards are not controlling in nature and when administered in a non-interfering and non-controlling context, they can enhance both intrinsic motivation and performance. In this experiment, participants will receive rewards in two types of salient versus non-salient performance-contingent pay and in controlling versus non-controlling contexts. Our two main hypotheses are that non-salient performance-contingent rewards given in a non-controlling context will improve both intrinsic motivation and performance. However, salient performance-contingent rewards given in a controlling context will undermine both intrinsic motivation and performance.

In the second and third studies, we narrowed our research to referral decision making in health care systems. We were interested in finding the effect of different motivational factors on the referral performance of general practitioners. In the second study, the effect of social relationships on referral behavior, we designed an experiment to examine how social relationships between a gatekeeper, such as a general practitioner and a specialists could affect their referral patterns under centralized and decentralized referral systems. The results provided support for our hypothesis among doctors who had high confidence in their referral decision making. High-confidence doctors referred significantly fewer patients under the close relationship condition which is characteristic of decentralized referral systems. This result has important implications for health care systems. In recent years, there has been a shift towards centralized referral systems in order to improve the efficiency of the referral systems. Our results

show that decentralized referral systems may encourage more referrals by high-confidence doctors due to the lack of social safeguards that can impede inappropriate referrals among general practitioners.

In the third study, the effect of fundholding and pay-for-performance payment schemes on referral patterns, we designed an experiment to see how different payment schemes would affect referral rate and referral appropriateness. While we could not find statistically significant results to support our hypotheses in this study, probably due to the small sample size, the results were in the direction that we predicted: both fundholding and pay-for-performance schemes decreased the referral rate in comparison with the fixed pay condition. In addition, pay-for-performance scheme produced a smaller average referral error in comparison with the other two payment schemes.

Task referral is an important part of many operational systems such as call centers. In future studies, we would like to examine the effect of gatekeeper-specialist relationship and payment schemes on referral rate and referral appropriateness in similar referral systems such as call centers. Based on the results of our experiment, we hypothesize that gatekeeper-specialist relationship will result in lower referral rate among high-confidence gatekeepers. In addition, we would like to test the effect of different payment methods on gatekeepers' referral rate and referral appropriateness in other referral systems.

One of the main reasons for not obtaining results that support our hypotheses in the second and third studies would be the use of family practice residents rather than family physicians as participants in these two studies. Family practice residents do not have the referral knowledge, experience, and confidence that family physicians gain during the years of their

practice. In future studies, we would like to test our hypotheses by having family physicians participate in our study. Using family physicians can give better insight into the effect of social and monetary motivators on referral rate and referral appropriateness in referral systems. In additin to experimental analysis, we would like to test our hypotheses empirically by using real-world data from clinics that have both centralized and decentralized referral systems.

References

1. Adams, S., J., Pryor, L., J., & Adams, S., L. (1994). Attraction and attention of high aptitude students in accounting: An exploratory longitudinal study. *Issues in Accounting Education*; **9(1)**: 45-58.
2. Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, **50:** 179-211.
3. Ajzen, I. & Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: Prentice-Hall.
4. Akbari, A., Mayhew, A., Al-Alawi, M., A., Grimshaw, J., Winkens, R., Glidewell, E., Pritchard, C., Thomas, R. & Fraser, C. (2008). Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane Database of Systematic Reviews*; Issue 1.
5. Amabile,T., M., Hennesey, B., A., & Grossman, B., S. (1986). Social influence on creativity: The effects of contracted-for-rewards. *Journal of Personality and Social Psychology*; **50(1)**: 14-23.
6. Amundson, G., Solberg, L., I., Reed, M., Martini, M., E., & Carlson, R. (2003). Paying for quality improvement: compliance with tobacco cessation guidelines. *Joint Commission Journal on Quality and Patient Safety*; **29(2)**: 59-65.
7. An, L., C., Bluhm, J., H., Foldes, S., S., Alesci, N., L., Klatt, C., M., Center, B., A., Nerssian, W., S., Larson, M., E., Ahluwalia, J., S., & Manley, M., W. (2008). A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. *Archives of Internal Medicine*; **168(18)**: 1993-1999.

8. Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *Review of Economic Studies*; **76**: 451-469.
9. Armour, B., S., Pitts, M., M., Maclean, R., Cangialose, C., Kishel, M., Imai, H., & Etchason, J. (2001). The effect of explicit financial incentives on physician behavior. *Archives of Internal Medicine*; **161(10)**:1261-1266.
10. Audit Commission, (1996). *What the doctor ordered: A study of GP fundholders in England and Wales.* HMSO, London.
11. Bachman, K., H., & Freeborn, D., K. (1999). HMO physicians' use of referrals. *Social Science & Medicine*; **48(4)**: 547–57.
12. Baiman, S. (1990). Agency research in managerial accounting: A second look. *Accounting, Organizations and Society*; **15(4)**: 341–371.
13. Bandura, A. (1997). *Self-efficacy: The Exercise of Control*, *Freeman*, New York, NY.
14. Bandura, A. (1991a). Social cognitive theory of self-regulation. *Organizational behavior and human decision processes, 50*(2), 248-287.
15. Bandura, A. (1991). Self-regulation of motivation through anticipatory and self-reactive mechanisms. In *Perspectives on motivation: Nebraska symposium on motivation* (Vol. 38, pp. 69-164).
16. Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efftcacy and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology,* 41, 586-598.

17. Baumeister, R., F. (1984). Choking under pressure: Self-Consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*; **46(3)**: 610–620.

18. Beaulieu, N., D., & Horrigan, D., R. (2005). Putting smart money to work for quality improvement. *Health Services Research*; **40(5)**:1318-1334.

19. Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology.* Vol. 6. New York: Academic Press.

20. Bendoly, E., Donohue, K., & Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management*; **24(6)**: 737-752.

21. Bhattacharyya, T., Freiberg, A., A., Mehta, P., Katz, J., N., Ferris, T. (2009). Measuring the report card: the validity of pay-for-performance metrics in orthopedic surgery. *Health Affairs (Millwood)*; **28(2)**:526-532.

22. Bindra, D. (1959). *Motivation: A Systematic Reinterpretation*. New York: Ronald.

23. Birkmeyer, J., D., Sharp, S., M., Finlasyon, S., R., G., Fisher, E., S., & Wennberg, J., E. (1998). Variation profiles of common surgical procedures. *Surgery*; **124(5)**: 917- 923.

24. Boggiano, A., K., Harackiewicz, J., M., Bessette, J., M., & Main, D., S. (1985). Increasing children's interest through performance-contingent reward. *Social Cognition*; **3(4)**: 400-411.

25. Boggiano, A., K., Ruble, D., N., & Pittman, T., S. (1982). The mastery hypothesis and the overjustification effect. *Social Cognition*; **1(1)**: 38-49.

26. Bonner, S., E. & Sprinkle, G., B. (2002). The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research. *Accounting, Organizations, and Society*; **27(4-5)**: 303-345.

27. Bonner, S., Hastie, E., R., Sprinkle, G., B., & Young, S., M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: implications for management accounting. *Journal of Management Accounting Research*; **12(1)**: 19-64.

28. Boyle, J., F. (1985). Regional variations in the use of medical services and the accountability of the profession. *JAMA-Journal of the American Medical Association*; **254(3)**: 407.

29. Bradlow, J., & Coulter, A. (1993). Effect of fundholding and indicative prescribing schemes on general practitioners' prescribing costs. *British Medical Journal*; **307(6913)**: 1186-1189.

30. Brennan, T., P., & Glover, J., A. (1980). An examination of the effect of extrinsic reinforcers on intrinsically motivated behavior: Experimental and theoretical. *Social Behavior and Personality*; **8(1)**: 27-32.

31. Brereton, L., & Vasoodaven, V. (2010). The impact of the NHS market: an overview of the literature. *CIVITAS: Institute for the Study of Civil Society*.

32. Brudevold, C., McGhee S., M., & Ho, L., M. (2000). Contract medicine arrangements in Hong Kong: an example of risk-bearing provider networks in an unregulated environment. *Social Science & Medicine*; **51(8)**: 1221-1229.

33. Bull, A., Silcocks, P., Start, R., & Kennedy, A. (1991) General practitioners, skin lesions and the new contract. *Journal of Public Health*; **14(3)**: 300-306.

34. Calder, B., J., & Staw, B., M. (1975a). Interaction of intrinsic and extrinsic motivation: some methodological notes. *Journal of Personality and Social Psychology*; **31(1)**: 76-80.

35. Calder, B., J., & Staw, B. M. (1975b). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology*; **31(4)**: 599-605.

36. Camerer, C., F., & Hogarth, R., M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*; **19(1-3)**: 7-42.

37. Camerer, C., F., Lowenstein, G., & Prelec, D. (2005). Neuroeconomics: how neuroscience can inform economics. *Journal of Economic Literature*; **43(1)**: 9-64.

38. Cameron, J., & Pierce, W., D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*; **64(3)**: 363-423.

39. Campbell, S., Reeves, D., Kontopantelis, E., Middleton, E., Sibbald, B., & Rolnad, M. (2007) Quality of primary care in England with the introduction of pay for performance; *New England Journal of Medicine*; **357**(2): 181-190.

40. Carlsen, F. & Grytten, J. (2000). Consumer satisfaction and supplier induced demand. *Journal of Health Economics*; **19(5)**: 731-753.

41. Chaix-Couturier, C., Durand-Zaleski, I., Jolly, D., & Durieux, P. (2000). Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *International Journal for Quality in Health Care*; **12(2)**: 133-142.

42. Chassin, M., R., Brook, R., H., Park, R., E., Keesey, J., Fink, A., Kosecoff, J., Kahn, K., Merrick, N., & Solomon, D., H. (1986). Variations in the use of medical and surgical

services by the Medicare population. *New England Journal of Medicine*; **314(5)**: 285-290.

43. Chib, V., S., De Martino, B., Shimojo, S., & O'Doherty, J., P. (2012). Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron***; 74(3)**: 582–594.
44. Christianson, J. B., Leatherman, S., & Sutherland, K. (2008). Lessons from evaluations of purchaser pay-for-performance programs a review of the evidence. *Medical Care Research and Review*; **65(6 suppl)**: 5S-35S.
45. Christensen, B., Sorensen, H., T., & Mabeck, C., E. (1989). Differences in referral rates from general practice. *Family Practice*; **6(1)**: 19-22.
46. Chung, R., S., Chernicoff, H., O., Nakao, K., A., Nickel, R., C., & Legorreta, A., P. (2003). A quality driven physician compensation model: four-year follow-up study. *Journal of Healthcare Quality*; **25(6)**: 31-37.
47. Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2nd ed., Erlbaum, Hillsdale, NJ.
48. Coleman, K., Reiter, K., L., & Fulwiler, D. (2007). The impact of pay-for-performance on diabetes care in a large network of community health centers. *Journal of Health Care for the Poor and Underserved*; **18(4)**:966-983.
49. Conrad, D., A, & Christianson, J., B. (2004) Penetrating the "black box": financial incentives for enhancing the quality of physician services. *Medical Care Research and Review*; **61(3)**: 37S-68S.

50. Conrad, D., A. & Perry, L. (2009). Quality-based financial incentives in health care: can we improve quality by paying for it? *Annual Review of Public Health*; **30**: 357-371.

51. Coulter, A. (1998). Managing demand at the interface between primary and secondary care. *British Medical Journal*; **316(7149)**: 1974-1976.

52. Coulter, A., & Bradlow, J. (1993). Effect of NHS reforms on general practitioners' referral patterns. *British Medical Journal*; **306**: 433-437.

53. Coulter, A., Seagroatt, V., & McPherson, K. (1990). Relation between general practices' outpatient referral rates and rates of elective admission to hospital. *British Medical Journal*; **301(6746)**: 273-276.

54. Crombie, D., L., & Fleming, D. (1988). General practitioners' referrals to hospital: the financial implications of variability. *Health Trends*; **20(2)**: 53-56.

55. Cromwell, J., & Mitchell, J., B. (1986). Physician-induced demand for surgery. *Journal of Health Economics*; **5(4)**: 293-313.

56. Cummins, R., O., Jarman, B., & White, P., M. (1981). Do general practitioners have different 'referral thresholds'? *British Medical Journal*; **282(6269)**: 1037-1039.

57. Cunningham, M., R. (1988a). Does happiness mean friendliness? Induced mood and heterosexual self-disclosure. *Personality and Social Psychology Bulletin*; **14(2)**: 283-297.

58. Cunningham, M., R. (1988b). What do you do when you're happy or blue? Mood, expectancies, and behavioral interest. *Motivation and Emotion*; **12(4)**: 309-331.

59. Curtin, K., Beckman, H., Pankow, G., Milillo, Y., & Greene, R., A. (2006). Return on investment in pay for performance: A diabetes case study. *Journal of Healthcare Management*; **51(6)**: 365-374.

60. Daniel, T., L., & Esser, J., K. (1980). Intrinsic motivation as influenced by rewards, task interest, and task structure. *Journal of Applied Psychology*; **65(5)**: 566-573.

61. Davidson, S., M., Manheim, L., M., Werner, S., M., Hohlen, M., M., Yudkowsky, B., K., & Fleming, G., V. (1992). Prepayment with office-based physicians in publicly funded programs: results from the Children's Medicaid Program. *Pediatrics*; **89(4)**: 761-767.

62. de Brantes, F., S., & D'Andrea, B., G. (2009). Physicians respond to pay-for-performance incentives: larger incentives yield greater participation. *American Journal of Managed Care*; **15(5)**: 313-318.

63. deCharms, R. (1968). *Personal Causation: The Internal Affective Determinants of Behavior*. New York: Academic Press.

64. Deci, E., L., & Ryan, R., M. (2000). The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry*; **11(4)**: 227-268.

65. Deci, E., L., Ryan, M., R., & Koestner, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*; **125(6)**: 627-668.

66. Deci, E., L., & Ryan, R., M. (1985). *Intrinsic Motivation and Self-determination in Human Behavior*. New York: Plenum.

67. Deci, E., L., & Ryan, R., M. (1980). The Empirical Exploration of Intrinsic Motivational Processes. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 13, 39-80). New York: Academic Press.

68. Deci, E., L. (1976). The hidden cost of rewards. *Organizational Dynamics*; **4**: 61-72

69. Deci, E., L. (1972a). The Effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance*; **8(2)**: 217-229.

70. Deci, E., L. (1972b). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology*; **22(1)**: 113-120.

71. Deci, E. (1971) Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*; **18(1)**: 105-115.

72. De Coster, C., McMillan, S., Brant, R., McGurran, J., & Noseworthy T. & Primary Care Panel of the Western Canada Waiting List Project. (2007). The Western Canada waiting list project: Development of a priority referral score for hip and knee arthroplasty. *Journal of Evaluation in Clinical Practice*; **13**, 192-197.

73. De Jaegher, K., & Jegers, M. (2000). A model of physician behavior with demand inducement. *Journal of Health Economics*; **19(2)**: 231-258.

74. Delnoij, D., M., J., & Spreeuwenberg, P., M., M. (1997). Variation in GP's referral rates to specialists in internal medicine. *European Journal of Public Health*; **7(4)**: 427-435.

75. De Marco, P., Cain, C., Lockwood, T., & Roland, M. (1993). How valuable is feedback of information on hospital referral patterns? Lessons from visits to 92 East Anglian practices. *BMJ: British Medical Journal*; **307**: 1465-1466.

76. Department of Health, (2004). Practice based commissioning. 15 December. http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENTID=4098564&chk=uBbP%2Bg.

77. Devlin, R. A., & Sarma, S. (2008). Do physician remuneration schemes matter? The case of Canadian family physicians. *Journal of Health Economics*; **27(5)**: 1168-1181.

78. Devlin, R., A., Sarma, S., & Hogg, W. (2006). Remunerating primary care physicians: Emerging directions and policy options for Canada. *Healthcare Quarterly*; **9(3)**: 34-42.

79. Dickinson, A., M. (1989). The detrimental effects of extrinsic reinforcement on "intrinsic motivation". *The Behavior Analyst*; **12(1)**: 1-15.

80. Dijk, N. M., & Sluis, E. (2008). To pool or not to pool in call centers. *Production and Operations Management*; **17(3)**: 296-305.

81. Dixon, J., & Glennerster, H. (1995). What do we know about fundholding in general practice? *British Medical Journal*; **311(7007)**: 727-730.

82. Dixon, J., Dinwoodie, M., Hodson, D., Dodd, S., Poltorak, T., Garrett, C., Rice, P., Doncaster, I., & Williams, M. (1994). Distribution of NHS funds between fundholding and non-fundholding practices. *British Medical Journal*; **309(6946)**: 30–34.

83. Donaldson, C., & Gerard, K. (1984) Paying general practitioners: Shedding light on the review of health services. *Journal of Royal College of General Practitioners*; **39(320)**:114-117.

84. Doran, T., Fullwood, C., Kontopantelis, E., & Reeves, D. (2008). Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework. *Lancet*; **372(9640)**:728-736.

85. Dowell, J., S., Snadden, D., & Dunbar, J., A. (1995). Changing to generic formulary: how fundholding practice reduced prescribing costs. *British Medical Journal*; **310(6978)**: 505-508.

86. Downing, A., Rudge, G., Cheng, Y., Tu, Y., K., Keen, J., Gilthorpe, M., S. (2007). Do the UK government's new Quality and Outcomes Framework (QOF) scores adequately measure primary care performance? A cross-sectional survey of routine healthcare data. *BMC Health Services Research*; **7(1)**: 166.

87. Dudley, R., A., Frolich, A., Robinowitz, D., L., Talavera, J., A., Broadhead, P., & Luft, H., S. (2004). *Strategies to support quality-based purchasing: A review of the evidence.* Rockville, MD.

88. Dusheiko, M., Gravelle, H., Jacobs, R., & Smith, P. (2006). The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment. *Journal of Health Economics*; **25(3)**: 449-78.

89. Easterbrook, J., A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*; **66(3)**: 183-201.

90. Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: reality or myth? *American Psychologist*; **51(11)**: 1153-1166.

91. Eisenhardt, K., M. (1989). Agency theory: an assessment and review. *Academy of Management Review*; **14(1)**: 57–74.

92. Ellis, R., & McGuire, T. (1986). Provider behaviour under prospective reimbursement. Cost sharing and supply. *Journal of Health Economics*; **5(2)**: 129-151.

93. Elwyn, G., J., & Stott, N., C., H. (1994) Avoidable referrals? Analysis of 170 consecutive referrals to secondary care. *British Medical Journal*; **309(6954)**: 576-578.

94. Emmanuel, J., & Walter, N. (1989). Referrals from general practice to hospital outpatient departments: a strategy for improvement. *BMJ: British Medical Journal*; **299(6701)**: 722-724.

95. Enzle, M., & Ross, J. (1977). Increasing and decreasing intrinsic interest with contingent rewards: A test of the cognitive evaluation theory. *Journal of Experimental Social Psychology*; **14(6)**: 588-597.

96. Evans, R. (1974). Supplier induced demand: some empirical evidence and implications. In: Perlman M., ed. *The economics of health and medical care*. New York: Stockton Press, 163-173.

97. Evans, E., O., & McBride, K. (1968). Hospital usage by a group practice. A trainer-trainee study. *Journal of the Royal College of General Practitioners*; **16(4)**: 294-306.

98. Fairbrother, G., Hanson, K., L., Butts, G., C., & Friedman, S. (2001). Comparison of preventive care in Medicaid managed care and Medicaid fee for service in institutions and private practices. *Ambulatory Pediatrics*; **1(6)**: 294-301.

99. Fairbrother, G., Hanson, K., L., Friedman, S., & Butts, G., C. (1999). The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health*; **89(2)**: 171-175.

100. Fertig, A., Roland**,** M., King, H., & Moore, T. (1993). Understanding variation in rates of referral among general practitioners: are inappropriate referrals important and would guidelines help to reduce rates? *BMJ: British Medical Journal*; **307(6917)**: 1467-1470.

101. Fiske, S., T., Gilbert, D., T., & Lindzey, G. (2010). *Handbook of Social Psychology*, Fifth Edition. John Wiley & Sons, Inc.

102. Flodgren, G., Eccles, M., P., Shepperd, S., Scott, A., Parmelli, E., & Beyer, F., R. (2011). An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Systemic Review*; **7**.

103. Forrest, C., B., Nutting, P., A., von Schrader, S., Rohde, C., & Starfield, B. (2006). Primary care physician specialty referral decision making: patient, physician, and health care system determinants. *Medical Decision Making*; **26(1)**: 76-85.

104. Forrest C., Nutting, B., Starfield B., & von Schrader S. (1992). Family physicians' referral decision. *Journal of Family Medicine*; **51(3)**: 215-22.

105. Forsyth, G, & Logan, R., F., L. (1968). *Gateway or dividing line? A study of hospital out-patients in the 1960s.* Oxford University Press.

106. Franks, P.,Williams, G., C., Zwanziger, J., Mooney, C., & Sorbero, M. (2000). Why do physicians vary so widely in their referral rates? *Journal of General Internal Medicine*; **15(3)**:163-168.

107. Frey, B., S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*; **15(5)**: 589-611.

108. Frolich, A., Talavera, J., A., Broadhead, P., & Dudley, R., A. (2007). A behavioral model of clinician responses to incentives to improve quality. *Health Policy*; **80(1)**: 179-193.

109. Furnham, A., & Argyle, M. (1998). *The Psychology of Money*. New York, NY: Routledge.

110. Gandhi, T., K., Sitting, D., F., Franklin, M., Sussman, A., J., Fairchild, D., G., & Bates, D., W. (2000). Communication breakdown in the outpatient referral process. *Journal of General Internal Medicine*; **15(9)**: 626-631.

111. Gilmore, A., S., Zhao, Y., X., Kang, N., Ryskina, K., L., Legorreta, A., P., Taira, D., A., & Chung, R., S. (2007). Patient outcomes and evidence-based medicine in a preferred provider organization setting: A six-year evaluation of a physician pay-for-performance program. *Health Services Research*; **42(6p1)**:2140-2159.

112. Giuffrida, A., Gravelle, H., & Roland, M. (1999). Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes. *BMJ: British Medical Journal*, **319**(**7202**): 94-98.

113. Glasziou, P., P., Mar, C., D., Harris, M., Scott, A., Scott, I., A., & Stockwell, A. (2012). When financial incentives do more good than harm: a checklist. *BMJ: British Medical Journal*; **345**: e5047.

114. Glickman, S., W., Ou, F., S., Delong, E., R., Roe, M., T., Lytle, B., L., Mulgund, J., Rumsfeld, J., S., Gibler, W., B., Ohman, E., M., Schulman, K., A., Peterson, E., D. (2007). Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA-Journal of the American Medical Association*; **297(21)**: 2373-2380.

115. Gneezy, U. & Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics*; **115(3)**: 791-810.

116. Goldmann, F. (1952). Methods of payment for physicians' services in medical care programs. *American Journal of Public Health*; **42(2)**: 134–141.

117. Gosden, T., Forland, F., Kristiansen, I. S., Sutton, M., Leese, B., Giuffrida, A., Sergison, M., & Pedersen, L. (2001). Impact of payment method on behaviour of primary care physicians: a systematic review; *Journal of Health Services Research & Policy*; **6(1)**: 44-55.

118. Gosden, T., Pedersen, L., & Torgerson, D. (1999). How should we pay doctors? A systematic review of salary payments and their effect on doctor behavior. *Quarterly Journal of Medicine*; **92**: 47-55.

119. Grace, J., F., & Armstrong, D. (1987). Referrals to hospital: perceptions of patients, general practitioners and consultants about necessity and suitability of referral. *Family Practice*; **4(3)**: 170-175.

120. Grady, K., E., Lemkau, J., P., Lee, N., R., & Caddell, C. (1997), Enhancing mammography referral in primary care. *Preventive Medicine*, **26(6)**: 791-800.

121. Greene, S., E. & Nash, D., B. (2009). Pay for performance: an overview of the literature. *American Journal of Medical Quality*; **24(2)**: 140-163.

122. Greene, R., A., Beckman, H., Chamberlain, J., Partridge, G., Miller, M., Burden, D., & Kerr, J. (2004). Increasing adherence to a community-based guideline for acute sinusitis through education, physician profiling, and financial incentives. *American Journal of Managed Care*; **10(10)**: 670-678.

123. Grossbart, S., R. (2006). What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Medical Care Research and Review*; **63(1)**: 29S-48S.

124. Grytten, J., & Sorensen, R., J. (2001). Type of contract and supplier-induced demand for primary physicians in Norway. *Journal of Health Economics*; **20(3)**: 379-394.

125. Grytten, J., Carlsen, F., & Sorensen, R. (1995). Supplier inducement in a public health care system. *Journal of Health Economics*; **14(2)**: 207-229.

126. Haines, A., & Armstrong, D. (1992). Developing referral guidelines. In: Roland M. O., Coulter A, eds. Hospital referrals. *Oxford: Oxford University Press*, 163-178.

127. Hair, J., F., Black, W., C., Babin, B., J., Anderson, R., E., & Tatham, R., L. (2010). *Multivariate data analysis* (Vol. 7). Upper Saddle River, NJ: Prentice Hall.

128. Hajjaj, F., M., Salek, M., S., Basra, M., K., A., & Finlay, A., Y. (2010). Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. *Journal of the Royal Society of Medicine*; **103(5)**: 178-187.

129. Hamner, W., C., & Foster L., W. (1975). Are intrinsic and extrinsic rewards additive: A test of Deci's cognitive evaluation theory of task motivation. *Organizational Behavior and Human Performance*; **14(3)**: 398-415.

130. Harackiewicz, J., M., & Manderlink, G. (1984). A process analysis of the effects of performance-contingent rewards in intrinsic motivation. *Journal of Experimental Social Psychology*, **20(6)**: 531-551.

131. Harackiewicz, J., M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: The effects of evaluation on intrinsic interest. *Journal of Personality and Social Psychology*; **47(2)**: 287-300.

132. Hausman, D. & LeGrand, J. (1999). Incentives and health policy: primary and secondary care in the British National Health Service. *Social Science and Medicine*, **49(10)**: 1299–1307.

133. Helfinger, C., A., & Northrup, D., A. (2000). What happens when capitated behavioural health comes to town? The transition from the Fort Bragg demonstration to a capitated managed behavioural health contract. *Journal of Behavioural Health Services and Research*; **27(4)**: 390-405.

134. Helliwell, P., S., & Wright, V. (1991). Referrals to rheumatology. *BMJ: British Medical Journal*; **302(6772)**: 304-305.

135. Herrin, J., Nicewander, D., & Ballard, D., J. (2008). The effect of health care system administrator pay-for-performance on quality of care. *Joint Commission Journal on Quality and Patient Safety*; **34(11)**: 646-654.

136. Heyman, J., & Ariely, D. (2004). Effort for payment: A tale of two markets. *Psychological Science*; **15(11)**: 787-793.

137. Hickson, B., Altemeier, A., & Perrin, M. (1987). Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. *Pediatrics*; **80(3)**: 344-50.

138. Higgins, E., T., & Trope, Y. (1990) Activity Engagement Theory: Implications of Multiply Identifiable Input for Intrinsic Motivation. In E. T. Higgins & R. M. Sorrentino

(Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (Vol. 2, 229-264). New York: Guilford.

139. Hillman, A., L., Ripley, K., Goldfarb, N., Weiner, J., Nuamah, I., & Lusk, E. (1999). The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics*; **104(4 Pt 1)**: 931-935.

140. Hillman, A., L., Ripley, K., Goldfarb, N., Nuamah, I., Weiner, J., & Lusk, E. (1998). Physician financial incentives and feedback: Failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health*; **88(11)**: 1699-1701.

141. Himmel, W., Kron, M., Thies-Zajonc, S., & Kochen, M. (1997). Changes in drug prescribing under the public health reform law- a survey of general practitioners' attitude in East and West Germany. *International Journal of Clinical Pharmacology and Therapeutics*; **35(4)**: 164-168.

142. Hippisley-Cox, J., Hardy, C., Pringle, M., Fielding, K., Carlisle, R., & Chilvers, C. (1997). The effect of deprivation on variations in general practitioners' referral rates: a cross sectional study of computerized data on new medical and surgical outpatient referrals in Nottinghamshire. *BMJ: British Medical Journal*; **314(7092)**: 1458-1461.

143. Holden, M., & Madore, O. (2002). *Remuneration of Primary Care Physicians*. Library of Parliament, Parliamentary Research Branch.

144. Holtgrave, D., R., Lawler, F., & Spann, S., J. (1990). Physician's risk attitude, laboratory usage, and referral decisions: the case of an academic family practice center. *Society for Medical Decision Making*; **11(25)**: 124-130.

145. Hughes, D., & Yule, B. (1992). The effect of per-item fees on the behaviour of general practitioners. *Journal of Health Economics*; **11(4)**: 413–437.

146. Hutchison, B., Abelson, J., & Lavis, J. (2001). Primary care in Canada: so much innovation, so little change. *Health Affairs*; **20(3)**: 116-131.

147. Hutchinson, J., M., & Foley R., N. (1999). Method of physician remuneration and rates of Antibiotic prescription. *Canadian Medical Association Journal*; **160(7)**: 1013-1017.

148. Hutchison, A. (1993). Explaining referral variation: GPs cannot afford to be uninterested in the topic. *BMJ: British Medical Journal*; **307(6917)**: 1439.

149. Jackson, D., N., Hourany, L., & Vidmar, N., J. (1978). A four-dimensional interpretation of risk taking. *Journal of Personality*; **40(3)**: 483-501.

150. Jenkins, G., D., Mitra, A., Gupta, N., & Shaw, J. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology*; **83(5)**: 777-787.

151. Jones, T., V. (1991). Validity of case simulations. *Journal of Clinical Epidemiology*; **44(10)**: 1120-1121.

152. Jones, T., V., Gerrity, M., S., & Earp, J. (1990). Written case simulations: Do they predict physicians' behaviour? *Journal of Clinical Epidemiology*; **43(8)**: 805-815.

153. Jones, D., T. (1987). A survey of hospital outpatient refrral rates, Wales. *BMJ: British Medical Journal*; **295(6600)**: 734-736.

154. Jorgenson, D., Dunnette, M., D., & Pritchard, R., D. (1973). Effects of the manipulation of a performance-reward contingency on behavior in a simulated work setting. *Journal of Applied Psychology*; **57(3)**: 271-280.

155. Kahn, C., N., Ault, T., Isenstein, H., Potetz, L., & Van Gelder, S. (2006). Snapshot of hospital quality reporting and pay-for-performance under Medicare. *Health Affairs*; **25(1)**: 148-162.

156. Kammerling, R., M., & Kinnear, A. (1996). The extent of the two tier service for fundholders. *BMJ: British Medical Journal*; **312(7043)**: 1399–1401.

157. Kane, R., L., Johnson, P., E., Town, R., J., & Butler, M. (2004). Economic incentives for preventive care. *Evidence Reports/Technology Assessments*; **101**: 1-7.

158. Kanfer, R., & Ackerman, P., L. (1996). A self-regulatory skills perspective to reducing cognitive interference. In I., G. Sarason, B., R. Sarason, & G., R. Pierce (Eds.), *Cognitive interference: Theories, methods, and findings* (pp. 153–171). Mahwah, NJ: Erlbaum.

159. Kankaanpää, A., T., Franck, J., K., & Tuominen, R., J. (2012). Variations in primary care physicians' sick leave prescribing practices. *European Journal of Public Health*; **22(1)**: 92-96.

160. Kappes, H., B., & Oettingen, G. (2011) Positive fantasies about idealized futures sap energy. *Journal of Experimental Social Psychology*; **47(4)**: 719-729.

161. Karve, A., M., Ou, F., S., Lytle, B., L., & Peterson, E., D. (2008). Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *American Heart Journal*; **155(3)**: 571-576.

162. Kerssens, J., J., & Groenewegen, P., P. (1990). Referrals to physiotherapy: the relation between the number of referrals, the indication for referral and the inclination to refer. *Social Science & Medicine*; **30(7)**: 797-804.

163. Kim, D., Markham, F., S., & Cangelosi, J., D. (2002) Why students pursue a business degree: A comparison of business majors across universities. *Journal of Education for Business*; **78(1)**: 28-32.

164. Klinger, E. (1996). The contents of thoughts: Interference as the downside of adaptive normal mechanisms in thought flow. In J., A. Singer, & P. Salovey (Eds.), At Play in the Fields of Consciousness: Essays in Honor of Jerome L. Singer (29−50). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

165. Klinger, E. (1990). *Daydreaming: Using waking fantasy and imagery for self-knowledge and creativity.* Los Angeles: Tarcher.

166. Krasnik, A., Groenewegen, P., P., Pedersen, P., A., von Scholten, P., Mooney, G., Gottschau, A., ... & Damsgaard, M., T. (1990). Changing remuneration systems: effects on activity in general practice. *BMJ: British Medical journal*; **300(6741)**: 1698-1701.

167. Kruglanski, A., W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review*; **82(6)**: 387-406.

168. Kruglanski, A., W., & Friedman, I., & Zeevi, G. (1971). The effects of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality*, **39(4)**: 606-617.

169. Langer, E., J., & Imber, L., G. (1979). When practice makes imperfect: The debilitating effects of overlearning, *Journal of Personality and Social Psychology*, **37(11)**: 2014-2024.

170. Langley, G., R., Minkin, S., & Till, J., E. (1997). Regional variation in nonmedical factors affecting family physicians' decisions about referral for consultation. *Canadian Medical Association Journal*; **157(3)**: 265-272.

171. Langley, G., R., Tritchler, D., L., Llewellyn-Thomas, H., A., & Till, J., E. (1991). Use of written cases to study factors associated with regional variations in referral rates. *Journal of Clinical Epidemiology*; **44(4)**: 391-402.

172. Larsen, D., L., Cannon, W., & Towner, S. (2003). Longitudinal assessment of a diabetes care management system in an integrated health network. Journal of Managed Care Pharmacy; **9(6)**:552-558.

173. Lazear, E., P. (2000). Performance pay and productivity, *American Economic Review*; **90(5)**: 1346-1361.

174. Lee, Y. & Lee, S., J. (2006). The competitiveness of information systems major: An analytic hierarchy process. *Journal of Information Systems Education*; **17(2)**: 211-221.

175. Lepper, M., R., Sagotsky, G., Dafoe, J., L., & Greene, D. (1982). Consequences of superfluous social constraints: effects on young children's social inferences and subsequent intrinsic interest. *Journal of Personality and Social Psychology*; **42(1)**: 51-65.

176. Lepper, M., R. (1981). Intrinsic and Extrinsic Motivation in Children: Detrimental Effects of Superfluous Social Controls. In W. A. Collins (Ed.), *Aspects of the Development of Competence: The Minnesota Symposium on Child Psychology* (Vol. 14, 155-214). Hillsdale, NJ: Erlbaum.

177. Lepper, M., R., Greene, D., & Nisbett, R., E. (1973). Undermining children's intrinsic Interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology***; 28(1)**: 129-137.

178. Lerner, C. & Claxton, K. (1994). Modeling the behaviour of general practitioners: A theoretical foundation for studies of fundholding. York: University of York discussion paper series (no. 116).

179. Levaggi, R. & Rochaix, L. (2003). Optimal Payment Schemes for Physicians. *Portuguese Economic Journal*; **2(2)**: 87-107.

180. Levin-Scherz, J., DeVita, N., & Timbie, J. (2006). Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS ® measures in an integrated delivery network. *Medical Care Research and Review*; **63(1)**: 14S-28S.

181. Lewin, K. (1938). *The Conceptual Representation and Measurement of Psychological Forces.* Durham, N. C.: Duke University Press.

182. Liberman, N., Trope, Y., & Stephan, E. (2007). Psychological distance. In E.T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (Vol. 2, pp. 353–383). New York, NY: Guilford Press.

183. Lindenauer, P., K., Remus, D., Roman, S., Rothberg, M., B., Benjamin, E., M., Ma, A., & Bratzler, D., W. (2007). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine*; **356(5)**: 486-496.

184. Locke, E., A., & Latham, G., P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*; **15(5)**: 265-268.

185. Locke, E., A., & Latham, G., P. (2002). Building a practically useful theory of goal setting and task performance: A 35-year odyssey. *American Psychologist*; **57(9)**: 705-717.

186. Locke, E., A., & Latham, G., P. (1990). *A Theory of Goal Setting and Task Performance.* Englewood Cliffs, NJ: Prentice-Hall.

187. Locke, E., A., Shaw, K., N., Saari, L., M., & Latham, G., P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*; **90(1)**: 125-152.

188. Lockyer, J., Hofmeister, M., Crutcher, R., Klein, D., & Fidler, H. (2007). International medical graduates: Learning for practice in Alberta, Canada. *Journal of Continuing Education in the Health Professions*; **27(3)**: 157-163.

189. Loveland, K., K., & Olley, J., G. (1979). The Effect of external reward on Interest and quality of task performance in children of high and low intrinsic motivation. *Child Development*; **50(4)**: 1207-1210.

190. Madeley, R., J., Evans, J., R., & Muir, B. (1990). The use of routine referral data in the development of clinical audit and management in North Lincolnshire. *Journal of Public Health Medicine*; **12(1)**: 22-27.

191. Malgwi, C., A., Howe, M., A., & Burnaby, P., A. (2005). Influences on students' choice of college majors. *Journal of Education for Business*; **80(5)**: 275-282.

192. Marinker, M., Wilkin, D., Metcalfe, D., H. (1988). Referral to hospital: can we do better? *BMJ: British Medical Journal*; **297(6646)**:461-464.

193. Matsaganis, M., & Glennerster, H. (1994). The threat of 'cream skimming' in the post-reform NHS. *Journal of Health Economics*; **13(1)**: 31–60.

194. Maxwell, M., Heaney, D., Howie, J., G., R., & Noble, S. (1993). General practice fundholding: observations on prescribing patterns and costs using the daily dose method. *BMJ: British Medical Journal*; **307(6913)**: 1190-1194.

195. McBride, D., Hardoon, S., Walters, K., Gilmour, S., & Raine, R. (2010). Explaining variation in referral from primary to secondary care: cohort study. *BMJ: British Medical Journal*; **341**.

196. McElduff, P., Lyratzopoulos, G., Edwards, R., Heller, R., F., Shekelle, P., & Roland, M. (2004) Will change in primary care improve health outcomes? Modelling the impact of financial incentives introduced to improve quality of care in the UK. *Quality and Safety in Health Care*; **13(3)**: 191-197.

197. McKinlay, J., B., Lin, T., Freund, K., & Moskowitz, M. (2002). *Journal of Health and Social Behavior*; **43(1)**: 92-106.

198. McPherson K., Strong, P., M., Jones, L., & Britton, B., J. (1985). Do cholecystectomy rates correlate with geographic variations in the prevalence of gallstones? *Journal of Epidemiology and Community Health*; **39(2)**: 179-182.

199. McGuire, T. (2000). Physician agency. In A. Cuyler and J. Newhouse, eds., *The Handbook of Health Economics.* Amsterdam: Elsevier Science.

200. McIntyre, C., W., Watson, D., Clark, L., A., & Cross, S., A. (1991). The effect of induced social interaction on positive and negative affect. *Bulletin of the Psychonomic Society*; **29(1)**: 67-70.

201. McIntyre, C., W., Watson, D., & Cunningham, A., C. (1990). The effects of social interaction, exercise, and test stress on positive and negative affect. *Bulletin of the Psychonomic Society*; **28(2)**: 141–143.

202. Mehrotra, A., Damberg, C., L., Sorbero, M., E., S., & Teleki, S., S. (2009). Pay for performance in the hospital setting: What is the state of the evidence? *American Journal of Medical Quality*; **24(1)**: 19-28.

203. Morgan, M., Jenkins, L., & Ridsdale, L. (2007). Patient pressure for referral for headache : a qualitative study of GPs' referral behaviour. *The British Journal of General Practice*; **57(534)**: 29-35.

204. Morgan, M. (1981). The overjustification effect: A developmental test of self-perception interpretations. *Journal of Personality and Social Psychology*; **40(5)**: 809-821.

205. Morrell, D., C., Gage, H., G., & Robinson, N., A. (1971). Referral to hospital by general practitioners. *The Journal of the Royal College of General Practitioners*; **21(103)**: 77-85.

206. Morrow, R., W., Gooding, A., D., & Clark, C. (1995). Improving physicians' preventive health care behavior through peer review and financial incentives. *Archives of Family Medicine*; **4**(2): 165-169.

207. Mullen, K., J., Frank, R., G., & Rosenthal, M., B. (2009). *Can you get what you pay for? Pay-for-performance and the quality of healthcare providers.* Cambridge, Massachusetts.

208. Nadler, E., S., Sims, S., Tyrance, P., H., D., G., Fairchild, D., G., Brenan T., A., & Bates, D., W. (1999). Does a year make a difference? Changes in physician satisfaction and perception in an increasingly capitated environment. *American Journal of Medicine*; **107(1)**: 38-44.

209. Nadler, D. A. (1979). The effects of feedback on task group behavior: A review of the experimental research. *Organizational Behavior and Human Performance*; **23(3)**: 309-338.

210. Nahra, T., A., Reiter, K., L., Hirth, R., A., Shermer, J., E., & Wheeler, J., R., C. (2006). Cost-effectiveness of hospital pay-for-performance incentives. *Medical Care Research and Review*; **63(1)**: 49S-72S.

211. Neter, J., & Wasserman, W. (1974). *Applied linear statistical models*. Irwin, Homewood, Illinois.

212. Newton, J., Hayes, V., & Hutchinson, A. (1991). Factors influencing general practitioner referral decisions. *Family Practice*; **8(4)**: 308-313.

213. Nightingale, S., D. (1988). Risk preference and admitting rates of emergency room physicians. *Medica Care*; **26(1)**: 84-87.

214. Nightingale, S., D., & Grant, M. (1988). Risk preference and decision making in critical care situations. *Chest*; **93(4)**:684-687.

215. Nightingale, S., D. (1987a). Risk preference and laboratory use. *Medical Decision Making*; **7(3)**: 168-172.

216. Nightingale, S., D. (1987b). Risk preference and laboratory test selection. *Journal of General Internal Medicine*; **2(1)**: 25-28.

217. Notzon, F., C., Placek, P., J., & Taffel, S., M. (1987). Comparisons of national cesarean-section rates. New England Journal of Medicine; **31(5)**: 386-389.

218. Oettingen, G., & Mayer, D. (2002). The motivating function of thinking about the future: Expectations versus fantasies. *Journal of Personality and Social Psychology*; **83(5)**: 1198-1212.

219. Oettingen, G., Pak, H., & Schnetter, K. (2001) Self-regulation of goal-setting: turning free fantasies about the future into binding goals. *Journal of Personality and Social Psychology*; **80(5)**: 736-753.

220. Oettingen, G., Honig, G., & Gollwitzer, P., M. (2000) Effective self-regulation of goal attainment. *International Journal of Educational Research*; **33(7)**: 705-732.

221. Oettingen, G. (2000) Expectancy effects on behavior depend on self-regulatory thought. *Social Cognition*; **18(2)**: 101-129.

222. O'Donnell, C., A. (2000). Variation in GP referral rates: what can we learn from the literature? *Family Practice*; **17(6)**: 462-471.

223. O'Malley, A. S. & Reschovsky, J., D. (2011). Referral and consultation communication between primary care and specialist physicians. *Archives of Internal Medicine*; **171(1)**: 56-65.

224. Pallak, S., R., Costomiris, S., Sroka, S., & Pittman, T., S. (1982). School experience, reward characteristics, and intrinsic motivation. *Child Development*; **53(5)**: 1382-1391.

225. Pearson, S., D., Schneider, E., C., Kleinman, K., P., Coltin, K., L., & Singer, J., A. (2008). The impact of pay-for-performance on health care quality in Massachusetts, 2001-2003. *Health Affairs*; **27(4)**:1167-1176.

226. Philips, J., S., & Freedman, S., M. (1985). Contingent pay and intrinsic task interest: moderating effects of work values. *Journal of Applied Psychology*; **70(2)**: 306-313.

227. Pilote, L., Califf, R., M., Sap, S., Miller, D., P., Mark, D., B., Weaver, D., Gore, J., M., Armstrong, P., W., Magnus Ohman, E., & Topol, E., J. (1995). Regional variation across the United States in the management of acute myocardial infarction. *New England Journal of Medicine*; **333(9)**: 565-572.

228. Pinder, C., C. (1976). Additivity versus nonadditivity of intrinsic and extrinsic incentives: implications for work motivation, performance, and attitudes. *Journal of Applied Psychology*; **61(6)**: 693-700.

229. Pittman, T., S., Cooper, E., E., & Smith, T., W. (1977). Attribution of causality and the overjustification effect. *Personality and Social Psychology Bulletin*, **3(2)**: 280-283.

230. Porter, L., W., & Lawler, E., E. (1968). *Managerial Attitudes and Performance.* Homewood, Illinois: Irwin-Dorsey.

231. Pritchard, R., D., Leonard, D., W., Von Bergen Jr., C., W., & Kirk, R., J. (1976). The effects of varying schedules of reinforcement on human task performance. *Organizational Behavior and Human Performance*; **16(2)**: 205–230.

232. Redfern, J., & Bowling, A. (2000). Efficiency of Care at the Primary-Secondary Interface: Variations with GP Fundholding. *Health & Place*, **6(1)**:15-23.

233. Reid, F., D., A., Cook, D., G., & Majeed, A. (1999). Explaining variation in hospital admission rates between general practices: cross sectional study. *BMJ: British Medical Journal*, **319** (**7202**): 98-103.

234. Rethans, J., J. (1991). Validity of case simulations. *Journal of Clinical Epidemiology*; **44(10)**: 1119-1120.

235. Reynolds, G., A., Chitnis, J., G., & Roland, M., O. (1991). General practitioner outpatient referrals: do good doctors refer more patients to hospital? *BMJ: British Medical Journal*; **302(6787)**: 1250-1252.

236. Rice, T., H. (1983). The impact of changing Medicare reimbursmenet rates on physician-induced demand. *Medical Care*; **21(8)**: 803-815.

237. Robinson, J., C. (2001). Theory and practice in the design of physician payment incentives. *The Milbank Quarterly*; **79(2)**: 149-177.

238. Roland, M., O. (1992). Measuring appropriateness of hospital referrals. In: Roland M. O., Coulter A., eds. *Hospital referrals.* Oxford: Oxford University Press; 136-149.

239. Roland, M., O., Porter, R., Matthews, J., G., Redden, J., F., Simonds, G., W., & Bewley, B. (1991). Improving care: a study of orthopedic outpatient referrals. *BMJ: British Medical Journal*; **302(6785)**: 1124-1128.

240. Roland, M., O., Bartholomew, J., Morrell, D., C., McDermott, A., & Paul, E. (1990). Understanding hospital referral rates: a user's guide. *BMJ: British Medical Journal*; **301(6743)**: 98-102.

241. Rosen, B. (1989). Professional reimbursement and professional behaviour: emerging issues and research challenges. *Social Science & Medicine*; **29(3)**: 455-462.

242. Rosenthal, M., B., de Brantes, F., S., Sinaiko, A., D., Frankel, M., Robbins, R., D., & Young, S. (2008). Bridges to Excellence - Recognizing High-Quality Care: Analysis of Physician Quality and Resource Use. *American Journal of Managed Care*; **14(10)**: 670-677.

243. Rosenthal, M., B., & Frank, R., G. (2006). What is the empirical basis for paying for quality in health care? *Medical Care Research and Review*; **63(2)**: 135-157.

244. Rosenthal, M., B., Frank, R., G., Li, Z., H., & Epstein, A., M. (2005). Early experience with pay-for performance-From concept to practice. *JAMA*: *Journal of the American Medical Association*; **294(14)**: 1788-1793.

245. Rothkopf, M. H., & Rech, P. (1987). Perspectives on queues: Combining queues is not always beneficial. *Operations Research*; **35(6)**: 906-909.

246. Rummel, A., & Feinberg, R. (1988). Cognitive evaluation theory: A meta-analytic review of the literature. *Social Behavior and Personality*; **16(2)**: 147-164.

247. Ryan, A., M. (2009). Effects of the premier hospital quality incentive demonstration on Medicare patient mortality and cost. *Health Services Research*; **44(3)**: 821-842.

248. Ryan-Nicholls, K., D., & Racher, F., E. (2004) Investigating the health of rural communities: toward framework development. *Rural and Remote Health*; **4(1)**: 244.

249. Ryan R., M., & Deci E., L. (2000). Intrinsic and extrinsic motivation: classic definitions and new directions. *Contemporary Educational Psychology*; **25(1)**: 54-67.

250. Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of personality and social psychology*; **43(3)**: 450.

251. Ryan, R., M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*; **45(4)**: 736-750.

252. Sabatino, S., A., Habarta, N., Baron, R., C., Coates, R., J., Rimer, B., K., Kerner, J., Coughlin, S., S., Kalra, G., P., & Chattopadhyay, S. (2008). Interventions to increase

recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers - Systematic reviews of provider assessment and feedback and provider incentives. *American Journal of Preventive Medicine*; **35(1 suppl)**: S67-S74.

253. Salancik, G., R. (1975). Interaction effects of performance and money on self-perception of intrinsic motivation. *Organizational Behavior and Human Performance*; **13(3)**: 339-351.

254. Salize, H., J., Merkel, S., Reinhard, I., Twardella, D., Mann, K., & Brenner, H. (2009). Cost effective primary care-based strategies to improve smoking cessation: more value for money. *Archives of Internal Medicine*, **169(3)**: 230-235.

255. Samanta, A., & Roy, S. (1988). Referrals from general practice to a rheumatology clinic. *British Journal of Rheumatology*; **27(1)**: 74-76.

256. Sarma, S., Devlin, R., A., Belhadji, B., & Thind, A. (2010). Does the way physicians are paid influence the way they practice? The case of Canadian family physicians' work activity. *Health Policy*; **98(2-3)**: 203-217.

257. Schatz, M. (2008). Does pay-for-performance influence the quality of care? *Current Opinion in Allergy and Clinical Immunology*; **8(3)**: 213-221.

258. Scott, A., Sivey, P., Ait Ouakrim, D., Willenberg, L., Naccarella, L., Furler, J., et al. (2011). The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Systematic Review*; **9**.

259. Scott, S., & Hall, J. (1995) Evaluating the effects of GP remuneration: problems and prospects. *Health Policy*; **31(3)**: 183-195.

260. Seligman, M., E., P. (1975). *Helplessness: On Depression, Development, and Death*. San Fransisco: Freeman.

261. Sibley, L. & Weiner, J. (2011). An evaluation of access to health care services along the rural-urban continuum in Canada. *BMC Health Services Research*; **11(1)**: 20.

262. Skedgel, C. (1996). Alternatives for physician payment. *Population Health Research Unit*; Dalhousie University, Halifax.

263. Skinner, B., F. (1953). *Science and human behavior*. New York: Macmillan.

264. Skinner, B., F. (1945). The operational analysis of psychological terms. *Psychological Review*; **52(2)**, 270-277.

265. Sladden, M., J. & Graham-Brown, R., A. (1989). How many GP referrals to dermatology outpatients are really necessary? *Journal of the Royal Society of Medicine*; **82(6)**: 347-348.

266. Smith, D. R., & Whitt, W. (1981). Resource sharing for efficiency in traffic systems. *Bell System Technical Journal*; **60(1)**: 39-55.

267. Slovic, P., & Lichtenstein, S. (1983). Preference reversal: a broader perspective. *American Economic Review*; **73(4)**: 596-605.

268. Sorensen, R., J. & Grytten, J. (2000). Contract design for primary care physicians: physician location and practice behavior in small communities. *Health Care Management Science*; **3(2)**: 151-157.

269. Srirangalingam, U., Sahathevan, S., K., Lasker, S., S., & Chowdhury, T., A. (2006). Changing pattern of referral to a diabetes clinic following implementation of the new UK GP contract. *British Journal of General Practice*; **56(529)**: 624-626.

270. Stearns, S., Wolfe, B., L., Kindig, D., A. (1992) Physician responses to fee-for-service and capitation payment. *Inquiry*; **29(4)**: 416-425.

271. Steel, P. & MacDonnell R. (2012). When rewards go wrong: A tale of five motivational misdirects. *Performance Improvement*; **51(8)**: 19-25.

272. Surender, R., Bradlow, J., Coulter, A., Doll, H., & Stewart Brown, S. (1995). Prospective study of trends in referral patterns in fundholding and non-fundholding practices in the oxford region, 1990-4. *BMJ: British Medical Journal*; **311(7014)**: 1205-1208.

273. Swann, W., B., & Pittman, T., S. (1977). Initiating play activity of children: The moderating influence of verbal cues on intrinsic motivation. *Child Development*, **48(3)**: 1128-1132.

274. Tahrani, A., A., McCarthy, M., Godson, J., Taylor, S., Slater, H., Capps, N., Moulik, P., &Macleod, A., F. (2007). Diabetes care and the new GMS contract: the evidence for a whole county. *British Journal of General Practice*, **57(539)**: 483-485.

275. Tang, S., H., & Hall, V., C. (1995). The Overjustification Effect: A Meta-Analysis. *Applied Cognitive Psychology*; **9(5)**: 365-404.

276. Tarlov, A., R., Ware, J., E., Greenfield, S., Nelson, E., C., Perrin E., & Zubkoff, M. (1989). The medical outcomes study: an application of methods for monitoring the results of medical care. *Journal of the American Medical Association* **262**: 925-30.

277. Tolman, E., C. (1932). *Purposive Behavior in Animals and Men.* New York: Century.

278. Trope, Y. & Liberman, N. (2003). Temporal construal. *Psychological Review;* **110(3)**: 403-421.

279. Tu, K., Cauch-Dudeck, K., & Chen, Z. (2009). Comparison of primary care physician payment models in the management of hypertension. *Canadian Family Physician*, **55(7)**: 719-727.

280. Tussing, A., D., & Wojtowycz, M., A. (1986). Physician induced demand by Irish GPs. *Social Science & Medicine*; **23(9)**: 851-860.

281. Twardella, D., & Brenner, H. (2007). Effects of practitioner education, practitioner payment and reimbursement of patients' drug costs on smoking cessation in primary care: a cluster randomised trial. *Tobacco Control*; **16(1)**:15-21.

282. Vaghela, P., Ashworth, M., Schofield. P., & Gulliford, M., C. (2008). Population intermediate outcomes of diabetes under pay for performance incentives in England from 2004 to 2008. *Diabetes Care*; **32(3)**: 427-429.

283. Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., & Sermeus, W. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research*; **10(1)**: 247.

284. Vehvilainen, A., T., Kumpusalo, E., A., Voutilainen, S., O., & Takala, J., K. (1996). Does the doctors' professional experience reduce referral rates? Evidence from the Finnish referral study. *Scandinavian Journal of Primary Health Care*; **14(1)**: 13-20.

285. Verhaak, P., F., M. (1993). Analysis of referrals of mental health problems by general practitioners. *BMJ: British Journal of General Practice*; **43(370)**: 203-208.

286. Vroom, V. (1964). *Work and Motivation.* New York, NY: John Wiley.

287. Watson, D., Clark, L., A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*; **54(6)**, 1063-1070.

288. Weber, E., U. (1988). A descriptive measure of risk. *Acta Psychologica*; **69(2)**: 185-203.

289. Weber, V., Bloom, F., Pierdon, S., & Wood, C. (2008). Employing the electronic health record to improve diabetes care: A multifaceted intervention in an integrated delivery system. *Journal of General Internal Medicine*; **23(4)**: 379-382.

290. Weinberg, R., S., & Jackson, A. (1979). Competition and extrinsic rewards: Effects on intrinsic motivation and attribution. *Research Quarterly*; **50(4)**: 494-502.

291. Weiner, M., J. (1980). The effect of incentive and control over outcomes upon intrinsic motivation and performance. *Journal of Social Psychology*; **112(2)**: 247-254.

292. Wennberg, J. (1986). Which rate is right? *New England Journal of Medicine*; **314(5)**: 310-311.

293. White, R., W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*; **66(5)**: 297-333.

294. Whynes, D., K., Baines, D., L., & Tolley, K., H. (1995). GP fundholding and the costs of prescribing. *Journal of Public Health Medicine*; **17(3)**: 323-329.

295. Wiersma, U., J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology*; **65(2)**: 110-114.

296. Wilkin, D. (1992). Patterns of referral: explaining variation. *Roland, M & Coulter, A., editor(s). Hospital Referrals*. Oxford: Oxford University Press.

297. Wilkin, D., Metcalfe, D., H., & Marinker, M. (1989) The meaning of information on GP referral rates to hospitals. *Community Medicine*; **11(1)**: 65-70.

298. Wilkin, D., & Smith, A. (1987a) Explaining variation in general practitioner's referral to hospital. *Family Practice*; **4(3)**: 160-169.

299. Wilkin, D., & Smith, A., G. (1987b). Variation in general practitioners' referral rates to consultants. *The Journal of the Royal College of General Practitioners*; **37(301)**: 350-353.

300. Williams, B., W. (1980). Reinforcement, behavior constraints, and the overjustification effect. *Journal of Personality and Social Psychology*; **39(4)**: 599-614.

301. Williams, T., F., White, K., L, Andrews, L., P, Diamond, E., Greenberg, B., G., Hamrick, A., A., & Hunter, E., A. (1960). Patient referral to a university clinic: patterns in a rural state. *American Journal of Public Health and the Nation's Health*; **50(10)**: 1493-1507.

302. Wilson, R., P., H., Hatcher, J., Barton, S., & Walley, T. (1996). Influence of practice characteristics on prescribing in fundholding and non-fundholding general practices: an observational study. *BMJ: British Medical Journal*; **313(7057)**: 595-599.

303. Wimperis, B. & Farr, J. (1979). The effects of task content and reward contingency upon task performance and satisfaction. *Journal of Applied Social Psychology*; **9(3)**: 229-249.

304. Woodward, R., & Warren-Bolton, F. (1984). Considering the effects of financial incentives and professional ethics on 'appropriate' medical care. *Journal of Health Economics*; **3(3)**: 223–237.

305. Wright, H., J. (1968). General practice in South West England. *The Journal of the Royal College of General Practitioners*; **15(4)**: 302.

306. Yerkes, R., M., & Dodson, J., D. (1908). The relationship of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology of Psychology*; **18**: 459–482.

307. Yett, D., Der. W., Ernst, R., & Hay, J. (1983). Physician pricing and health insurance reimbursement. *Health Care Financing Review*; **5(2)**: 69–80.

308. Young, G., J., Meterko, M., Beckman, H., Baker, E., White, B., Sautter, K., M., Greene, R., Curtin, K., Bokhour, B., G., Berlowitz, D., & Burgess, J., F. (2007). Effects of paying physicians based on their relative performance for quality. *Journal of General Internal Medicine*; **22(6)**: 872-876.

309. Zelizer, V. (1994). *The Social Meaning of Money*. New York, NY: Basic Books.

310. Zhang, W. (2007) Why IS? Understanding undergraduate students' intentions to choose an information systems major. *Journal of Information Systems Education*; **18(4)**: 447-458.

APPENDICIES

APPENDIX A: FIRST STUDY'S SECTIONS

A.1. Preliminary Questionnaire

*[*All the sentences within brackets and in Italic format were removed in the study questionnaires*.]

Section 1: Preliminary Questions

Number: _____

Major: _____

1. Please indicate how interested you are in solving general mathematics problems on a scale from 1 to 7.

   Not interested at all 1 2 3 4 5 6 7 Very much interested

2. Please indicate how interested you are in answering English grammar questions on a scale from 1 to 7.

   Not interested at all 1 2 3 4 5 6 7 Very much interested

3. Which of the two following activities is more enjoyable to you (please select one):
   a. Solving mathematics problems
   b. Correcting English sentences grammatically

4. Please rate your level of interest in mathematics as a ***major*** on a scale from 1 to 7.

   No interest at all 1 2 3 4 5 6 7 Very high interest

5. Please rate your level of interest in English as a *major* on a scale from 1 to 7.

   No interest at all 1 2 3 4 5 6 7 Very high interest

6. Please rate how interesting mathematics is to you *in general* on scale from 1 to 7.

   Not interesting at all 1 2 3 4 5 6 7 Very interesting

7. Please rate how interesting English is to you *in general* on a scale from 1 to 7.

   Not interesting at all 1 2 3 4 5 6 7 Very interesting

8. Please select the reasons for choosing a literature-related or language-related field of study. *You can check as many items as you like.*

   a. I am interested in English language.
   b. My parents pushed me to choose this field of study.
   c. My friends and relatives suggested this field to me.
   d. There are many job opportunities for this field.
   e. Other

9. If you selected other, please indicate your reasons:

---

10. After finishing this section, you will be asked to take an English test which includes some *general* English questions. Please indicate how *confident* you are that you will perform well in this test. [*for students who took English test*]

    Not at all confident 1 2 3 4 5 6 7 Very confident

11. After finishing this section, you will be asked to take a mathematics test which includes some ***general*** mathematics questions. Please indicate how *confident* you are that you will perform well in this test. [*for students who took math test*]

    Not at all confident 1 2 3 4 5 6 7 Very confident

12. You are about to take a multiple choice test. Please rate your level of stress on a scale from 1 to 7.

    No stress at all 1 2 3 4 5 6 7 A great deal of stress

13. You can earn a considerable amount of money based on your performance in the test, how much do you think this amount will be useful to you? [*for piece-rate treatments only*]

    Not at all 1 2 3 4 5 6 7 Very much

A.2. Section 2: Math/English Test

A.2.1. Mathematics Test

[*For students in the mathematics test treatments*]

This section includes 30 general mathematics problems. You have 35 minutes to complete this part. **Note that you will earn $0.50 per correct answer** [*for piece-rate treatments only*]. Solve the problems and indicate the best of the answer choices given.

[*Mathematics test appears here.*]

A.2.2. English Test

[*For students in the English test treatments*]

This section includes 30 general English language questions. You have 35 minutes to complete this part. **Note that you will earn $0.50 for each correct answer** [*for piece-rate treatments only*] Read each question carefully and indicate the best of the answer choices given.

[*English test appears here.*]

A.3. After-the-test Questionnaire

Number: _____

Part 1

For each of the following statements, please indicate how true it is for you.

1. I felt **happy** while I was taking the test.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

2. I put a lot of effort in answering the questions in the test.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

3. I felt **upset** while I was taking the test.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

4. I devoted much energy to get the test done.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

5. I did NOT put much energy into this test.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

6. I felt **cheerful** while I was taking the test.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

7. I strived hard to become successful in the test.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

8. I enjoyed taking the test very much.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

9. I thought this was a boring test.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

10. Please estimate, on a scale from 0 to 100, what <u>percentage</u> of your answers was based on pure guessing?

_____

Part 2

This section includes background questions. Please, choose the answers that best apply to your background.

1. Your university year (Please check one):

   ☐First year ☐Second year ☐Third year ☐Fourth year ☐More

2. Your GPA _____

3. Your **SAT** test scores:
   a. Critical Reading (Verbal) section _____
   b. Mathematics section _____

4. How many **Mathematics** courses have you completed in the university? _____

5. How many **English** courses have you completed totally in the university? _____
6. Your gender (Please check one):

   ☐ Male ☐ Female

7. Your age:_____

8. Your ethnicity (Please check one):

   ☐ Caucasian

   ☐ First Nations

   ☐ Biracial

   ☐ Black

   ☐ Asian

   ☐ Middle Eastern

   ☐ Other: Please Specify:_____

9. Please indicate your first language:_____

10. Please indicate your second language:_____

11. Please indicate your country of birth: _____

Thank you very much for your participation!

Note: The questions in the mathematics/English test section were selected from the book: "The Official Guide for GMAT® Review", 12th Edition Copyright © 2009 by the Graduate Management Admission Council®. All rights reserved.
Published by Wiley Publishing, Inc., Hoboken, New Jersey

Reprinted with permission of John Wiley & Sons, Inc.

APPENDIX B: SECOND STUDY'S SECTIONS

B.1. Survey Description

Survey Description

This study is conducted by Dr. Diane Bischak, Associate Professor at the University of Calgary and Rosa Hendijani, PhD candidate at the Haskayne School of Business, the University of Calgary. The purpose of this study is to understand the factors affecting priority setting in the referral process. Your participation in this study is voluntary and you may refuse to participate altogether or may choose to withdraw from the study at any time. The survey will take approximately 30 minutes to complete. It includes a set of patient case descriptions in a specific area of specialty. It also includes additional questions, consisting of general background and demographic data. A panel of specialists has reviewed the patient case descriptions. After completing this survey we will provide you with performance feedback that is an opportunity for you to improve your referral skills. Upon the survey closing date, you will receive an email which includes your performance summary as compared with the panel of specialists' opinions concerning each patient case description.

The collected data will be kept on a computer disk, stored in a secured location and accessible only by the researchers for research purposes, including the completion of the co-investigator's PhD dissertation and the publication of scientific papers. Participation is completely voluntary, anonymous and confidential. No one except the principal investigator and the co-investigator will be allowed to see any of the answers to the questions. There are no names collected and attached to the responses. Only group information will be summarized for any presentation or publication of results. In case you withdraw from this study, the data collected to the point of withdrawal will be retained and used for research purposes.

Your e-mail address is the only required piece of personal contact information. **We respect your privacy and will not use your email address for any purpose other than the awarding of prizes and/or providing feedback regarding your performance in the survey. Your contact information will not be provided to any third party organization in any form nor will it be associated with your answers to the survey questions.** You do not need to provide any personal information if you do not want to participate in a draw for an Apple iPad 3.

There are no foreseeable risks, harms, or inconveniences associated with your participation in this study. The only cost on your part is the time you will spend for participating in this survey. **You can choose to take part in a draw for an Apple iPad 3 (with a value of approximately $550) after completing this survey. The number of prizes will reflect the number of overall respondents.** If you agree with the terms of this survey, please press **Yes** to proceed to the next page.

B.2. Preliminary Questionnaire

Section 1: Preliminary Questions

14. Please rate your knowledge level in each of the following areas of specialty on a scale from 1 to 7:

    a. Respirology

Not at all knowledgeable 1 2 3 4 5 6 7 Very knowledgeable

    b. Rheumatology

Not at all knowledgeable 1 2 3 4 5 6 7 Very knowledgeable

15. Please rate your confidence in your ability to determine whether a patient of yours should be referred to a specialist in these areas:

    a. Respirology

Not at all confident 1 2 3 4 5 6 7 Very confident

    b. Rheumatology

Not at all confident 1 2 3 4 5 6 7 Very confident

16. Please select the item that best applies to you:

    a. I am a medical resident.

    b. I am a family practice resident.

    c. I am a primary care physician.

    d. Other (Please explain your current status in the box below.)

B.3. Case Test

*[*All the sentences within brackets and in Italic format were removed in the study questionnaires.*]

[*The case test included a set of 12 patient case descriptions. For each case, the participant answered three referral-related questions. The case descriptions were anonymous. They were in one of the two areas of Rheumatology or Respirology. The Rheumatology cases were developed by Dr. Tom Noseworthy. The Respirology cases were developed by Dr. Sachin Pendharkar. Here is a sample of the case descriptions and their related questions for both close relationship and centralized referral system treatments.*]

Treatment 1 [*Close Relationship Treatments*]

Section 2.

The following case describes one of your patients who you may consider referring to a rheumatologist. The rheumatologist to whom you may refer the patient is Dr. Lee. Having referred many of your patients to Dr. Lee over the past few years, you consider yourself to have a close professional relationship with Dr. Lee. You trust and feel comfortable with Dr. Lee as a specialist. Please read the case and answer the questions below.

[*The case appears here.*]

Please indicate how urgent it is for this patient to see Dr. Lee:

0	1	2	3	4	5	6	7	8	9	10
Not Urgent										Very Urgent

Please indicate whether you would refer the patient to Dr. Lee:

☐ Yes ☐ No

Please briefly explain the reasons for your referral or non-referral decision:

_____

Treatment 2 [*Centralized Referral System Treatments*]

Section 2.

The following case describes one of your patients who you may consider referring to a rheumatologist. The referral system to which you refer the patient is a **centralized referral system** in which the patients are referred to different rheumatologists through a centralized system that allocates the patient to the rheumatologist that has the next available time slot. Please read the case and answer the questions below.

[*The case appears here.*]

Please indicate how urgent it is for this patient to see a rheumatologist:

0 1 2 3 4 5 6 7 8 9 10
Not Urgent Very Urgent

Please indicate whether you would refer the patient to a rheumatologist:

☐ Yes ☐ No

Please briefly explain the reasons for your referral or non-referral decision:

_____

B.4. After-the-test Questionnaire

Please indicate your overall view of the patient case descriptions used in this study:

_____

Section 3

We would like to understand about characteristics that may affect your decision making. For each of the following statements, please indicate how true it is for you:

1. I felt **upset** while I was completing the survey.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

2. I felt **cheerful** while I was completing the survey.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

3. I am outspoken.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

4. I am concerned about getting hurt.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

5. I enjoy an element of physical danger.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

6. I would not speak out in public on unpopular issues unless I were fairly certain of my position.
   ☐ True ☐ False

7. I am careful about not hurting people's feelings.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

8. I am generally risk-averse when dealing with patients.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

9. If asked to choose between the two following therapies for your patient, which one would you select? [*Risk-related item*]

   ☐ Therapy A, which **definitely** gives the patient five more years of life than the average longevity
   ☐ Therapy B, which gives a 50:50 chance of zero or ten years of additional life for your patient

10. If asked to choose between the two following therapies for your patient, which one would you select? [*Loss aversion item*]

☐ Therapy A, which the patient definitely lives five years less than the average person

☐ Therapy B, which gives a 50:50 chance of losing zero or ten years of additional life

for the patient

Section 4

This section includes background questions. Please, choose the answers that best apply to your background.

12. Year(s) of residency program: _____

13. Type of location of your residency program:

    a. Urban

    b. Rural

14. Your gender (Please check one):

☐ Male ☐ Female

15. Your age group:

    a. 20-29

    b. 30-39

    c. 40-49

    d. 50-59

    e. 60-69

    f. Above 69

16. Your ethnicity (Please check one):

☐ Caucasian

☐ First Nations

☐ Biracial

☐ Black

☐ Asian

☐ Middle Eastern

☐ Other

☐ I prefer not to answer

17. Is English your first language?

☐ Yes ☐ No

Thank you very much for participating in this survey.

If you would like to take part in the draw for an Apple iPad3 and/or receive feedback regarding your performance, please provide your email address below. We will contact you through your email address if you are one of our contest winners.

Please type your email address here:

_____

Please check the items that apply to your preferences (You can select both items.):

☐ I would like to receive feedback regarding my performance.

☐ I would like to participate in the Apple iPad 3 contest.

Please **DO NOT** close the survey link. Press "**Next**" to submit your response.

Note: The cases used in this study are developed by the Western Canada Waiting List (WCWL) group and are provided to us by Dr. Tom Noseworthy. [*Rhuematology treatments only*]

Note: The cases used in this study are developed and provided to us by Dr. Sachin Pendharkar. [*Respirology treatments only*]

APPENDIX C: THIRD STUDY'S SECTIONS

*[*All the sentences within brackets and in Italic format were removed in the study questionnaires*.]

C.1. Survey Description

Survey Description [*Treatment 1: Fixed Payment*]

This study is conducted by Rosa Hendijani, PhD candidate and Dr. Diane Bischak, Associate Professor at the Haskayne School of Business, the University of Calgary. This research protocol is approved by the Conjoint Health Research Ethics Board (CHREB). The purpose of this research is to investigate and better understand the factors affecting priority setting in the patient referrals from general practitioners to specialists. **You will receive $25 as your participation fee in this survey.**

Your participation in this study is voluntary and you may refuse to participate altogether or may choose to withdraw from the study at any time. The survey will take approximately 30 minutes to complete. It includes a set of patient case descriptions in different areas of specialty. It also includes additional questions, consisting of general background and demographic data. A panel of specialists has reviewed the patient case descriptions. After completing this survey you will receive a report which includes your responses and the panel of specialists' responses concerning each patient case description.

The collected data will be kept on a computer disk, stored in a secured location and accessible only by the researchers for research purposes, including the completion of the co-investigator's PhD dissertation and the publication of scientific papers. Participation is completely voluntary, anonymous and confidential. No one except the principal investigator and the co-investigator will be allowed to see any of the answers to the questions. There are no names collected and attached to the responses. Only group information will be summarized for any presentation or publication of results. In case you withdraw from this study, the data collected to the point of withdrawal will be retained and used for research purposes.

Your e-mail and mailing addresses are the only required pieces of personal contact information. **We respect your privacy and will not use your addresses for any purposes other than providing your participation fee and/or sending survey reports. Your contact information will not be associated with your answers to the survey questions.** You do not need to provide any personal information if you do not want to receive performance feedback and participation fee.

There are no foreseeable risks, harms, or inconveniences associated with your participation in this study. The only cost on your part is the time you will spend for participating in this survey. If you agree with the terms of this survey, please press **Next** to proceed with the rest of the survey.

Survey Description [*Treatment 2: Fundholding Policy*]

[*Survey description for the fundholding treatment is the same as the one for the fixed payment treatment.*]

This study is conducted by Rosa Hendijani, PhD candidate and Dr. Diane Bischak, Associate Professor at the Haskayne School of Business, the University of Calgary. This research protocol is approved by the Conjoint Health Research Ethics Board (CHREB). The purpose of this research is to investigate and better understand the factors affecting priority setting in the patient referrals from general practitioners to specialists. **You will receive $25 as your participation fee in this survey.**

Your participation in this study is voluntary and you may refuse to participate altogether or may choose to withdraw from the study at any time. The survey will take approximately 30 minutes to complete. It includes a set of patient case descriptions in different areas of specialty. It also includes additional questions, consisting of general background and demographic data. A panel of specialists has reviewed the patient case descriptions. After completing this survey you will receive a report which includes your responses and the panel of specialists' responses concerning each patient case description.

The collected data will be kept on a computer disk, stored in a secured location and accessible only by the researchers for research purposes, including the completion of the co-investigator's PhD dissertation and the publication of scientific papers. Participation is completely voluntary, anonymous and confidential. No one except the principal investigator and the co-investigator will be allowed to see any of the answers to the questions. There are no names collected and attached to the responses. Only group information will be summarized for any presentation or publication of results. In case you withdraw from this study, the data collected to the point of withdrawal will be retained and used for research purposes.

Your e-mail and mailing addresses are the only required pieces of personal contact information. **We respect your privacy and will not use your addresses for any purposes other than providing your participation fee and/or sending survey reports. Your contact information will not be associated with your answers to the survey questions.** You do not need to provide any personal information if you do not want to receive performance feedback and participation fee.

There are no foreseeable risks, harms, or inconveniences associated with your participation in this study. The only cost on your part is the time you will spend for participating in this survey. If you agree with the terms of this survey, please press **Next** to proceed with the rest of the survey.

Survey Description [*Treatment 3: Pay-for-performance Treatment*]

This study is conducted by Rosa Hendijani, PhD candidate and Dr. Diane Bischak, Associate Professor at the Haskayne School of Business, the University of Calgary. This research protocol is approved by the Conjoint Health Research Ethics Board (CHREB). The purpose of this research is to investigate and better understand the factors affecting priority setting in the patient referrals from general practitioners to specialists. **You will receive $25 as your participation fee in this survey. In addition to the $25 participation fee, you will receive an additional amount of a maximum of $25. The amount of this additional payment depends on your performance in the survey.**

Your participation in this study is voluntary and you may refuse to participate altogether or may choose to withdraw from the study at any time. The survey will take approximately 30 minutes to complete. It includes a set of patient case descriptions in different areas of specialty. It also includes additional questions, consisting of general background and demographic data. A panel of specialists has reviewed the patient case descriptions. After completing this survey you will receive a report which includes your responses and the panel of specialists' responses concerning each patient case description.

The collected data will be kept on a computer disk, stored in a secured location and accessible only by the researchers for research purposes, including the completion of the co-investigator's PhD dissertation and the publication of scientific papers. Participation is completely voluntary, anonymous and confidential. No one except the principal investigator and the co-investigator will be allowed to see any of the answers to the questions. There are no names collected and attached to the responses. Only group information will be summarized for any presentation or publication of results. In case you withdraw from this study, the data collected to the point of withdrawal will be retained and used for research purposes.

Your e-mail and mailing addresses are the only required pieces of personal contact information. **We respect your privacy and will not use your addresses for any purposes other than providing your participation fee and/or sending survey reports. Your contact information will not be associated with your answers to the survey questions.** You do not need to provide any personal information if you do not want to receive performance feedback and participation fee.

There are no foreseeable risks, harms, or inconveniences associated with your participation in this study. The only cost on your part is the time you will spend for participating in this survey.

If you agree with the terms of this survey, please press **Next** to proceed with the rest of the survey.

C.2. Preliminary Questionnaire

Section 1: Preliminary Questions

17. Please rate your knowledge level in each of the following areas of specialty on a scale from 1 to 7:

    a. Respirology

Not at all knowledgeable 1 2 3 4 5 6 7 Very knowledgeable

    b. Rheumatology

Not at all knowledgeable 1 2 3 4 5 6 7 Very knowledgeable

18. Please rate your confidence in your ability to determine whether a patient of yours should be referred to a specialist in these areas:

    a. Respirology

Not at all confident 1 2 3 4 5 6 7 Very confident

    b. Rheumatology

Not at all confident 1 2 3 4 5 6 7 Very confident

19. Please select the item that best applies to you:

    a. I am a medical resident.

    b. I am a family practice resident.

    c. I am a primary care physician.

    d. Other (Please explain your current status in the box below.)

C.3. Case Test

[*Part2: Case Test*]

[*The case test includes 12 patient case descriptions. For each case, the participant should answer three referral-related questions. The case descriptions are anonymous. They are in two areas of Rheumatology and Respirology. Here is a sample of the case descriptions and their related questions.*]

Treatment 1 [*Fixed Payment Treatment*]

Section 2.

The following case describes one of your patients who you may consider referring to a specialist. Please read the case and answer the questions below.

[*The case appears here.*]

- Please indicate how urgent it is for this patient to see a specialist:

0 1 2 3 4 5 6 7 8 9 10
Not Urgent Very Urgent

- Please indicate whether you would refer the patient to a specialist:

☐ Yes ☐ No

- Please briefly explain the reasons for your referral or non-referral decision:

_____

Treatment 2 [*Fundholding Treatment*]

Please read the following explanation carefully before you proceed with the rest of the study.

In this section you will be presented with a set of cases that describe your patients. You may consider referring some of them to a specialist. Assume that you have a budget of $1000 for the referral of these patients. Any time you refer a patient to a specialist, you will pay $100 out of

the budget. If any of your budget is left over at the end, 5% of the remainder will be given to your practice as a medical resident. Please read the cases and answer the related questions.

[*The case appears here.*]

- Please indicate how urgent it is for this patient to see a specialist:

0 1 2 3 4 5 6 7 8 9 10
Not Urgent Very Urgent

- Please indicate whether you would refer the patient to a specialist:

☐ Yes ☐ No

- Please briefly explain the reasons for your referral or non-referral decision:

_____

Treatment 3 [*Pay for Performance Treatment*]

Please read the following explanation carefully before you proceed with the rest of the study.

In this section you will be presented with a set of cases that describe your patients. You may consider referring some of them to a specialist. In addition to your participation fee, you can receive an extra amount of a maximum of $25 based on how well your ratings match the average ratings of a panel of specialists who have reviewed the same cases. Please read each case carefully and answer the related questions.

[*The case appears here.*]

- Please indicate how urgent it is for this patient to see a specialist:

0 1 2 3 4 5 6 7 8 9 10
Not Urgent Very Urgent

- Please indicate whether you would refer the patient to a specialist:

☐ Yes ☐ No

- Please briefly explain the reasons for your referral or non-referral decision:

_____

C.4. After-the-test Questionnaire

Please indicate your overall view of the patient case descriptions used in this study:

_____

Section 3

We would like to understand about characteristics that may affect your decision making. For each of the following statements, please indicate how true it is for you:

11. I felt **upset** while I was completing the survey.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

12. I felt **cheerful** while I was completing the survey.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

13. I am outspoken.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

14. I am concerned about getting hurt.

1	2	3	4	5	6	7
Not at all true			Somewhat True			Very true

15. I enjoy an element of physical danger.

2 2 3 4 5 6 7

Not at all true Somewhat True Very true

16. I would not speak out in public on unpopular issues unless I were fairly certain of my position.
☐ True ☐ False

17. I am careful about not hurting people's feelings.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

18. I am generally risk-averse when dealing with patients.

1 2 3 4 5 6 7

Not at all true Somewhat True Very true

19. If asked to choose between the two following therapies for your patient, which one would you select?

☐ Therapy A, which **definitely** gives the patient five more years of life than the average longevity
☐ Therapy B, which gives a 50:50 chance of zero or ten years of additional life for your patient

20. If asked to choose between the two following therapies for your patient, which one would you select? [*Loss aversion item*]
☐ Therapy A, which the patient definitely lives five years less than the average person

☐ Therapy B, which gives a 50:50 chance of losing zero or ten years of additional life for the patient

Section 4

This section includes background questions. Please, choose the answers that best apply to your background.

18. Year(s) of residency program: _____

19. Type of location of your residency program:

    a. Urban

    b. Rural

20. Your gender (Please check one):

☐ Male ☐ Female

21. Your age group:

    a. 20-29

    b. 30-39

    c. 40-49

    d. 50-59

    e. 60-69

    f. Above 69

22. Your ethnicity (Please check one):

☐ Caucasian

☐ First Nations

☐ Biracial

☐ Black

☐ Asian

☐ Middle Eastern

☐ Other

☐ I prefer not to answer

23. Is English your first language?

☐ Yes ☐ No

If you would like to receive your participation fee (plus any additional money that you have earned in the survey) [*The part in brackets is only for the pay for performance treatment.*], and/or a report comprised of your responses and a panel of specialists' responses to the patient case descriptions, please provide your email and mailing addresses below. We will send you a cheque to your mailing address

Email Address:

Mailing Address:

Please check the items that apply to your preferences (You can select both items.):

☐ I would like to receive feedback regarding my performance.

☐ I would like to receive my participation fee (plus any additional amount I have earned in the survey) [*The part in brackets is only for the pay for performance treatment.*].

Thank you very much for participating in this survey.

Please **DO NOT** close the survey link. Press "**Next**" to submit your response.

Note: The rheumatology cases used in this study are developed by the Western Canada Waiting List (WCWL) group and are provided to us by Dr. Tom Noseworthy. The respirology cases used in this study are developed and provided to us by Dr. Sachin Pendharkar.