A Study on Validity of Mongolian Civil Service Entrance Examination for Job Performance

직무성과를 위한 몽골 공무원 채용시험의 타당성에 관한 연구

August 2014

Graduate School of Public Administration

Seoul National University

Public Administration Major

Amartuvshin Amarmend

A Study on Validity of Mongolian Civil Service Entrance Examination for Job Performance

Eom, SeokJin

Submitting a master's thesis of Public Administration

April 2014

Graduate School of Public Administration

Seoul National University

Public Administration Major

Amartuvshin Amarmend

Confirming the master's thesis written by

Amartuvshin Amarmend

June 2014

Chair   <u>Kim, SoonEun</u> (Seal)

Vice Chair   <u>Lee, SooYoung</u> (Seal)

Examiner   <u>Eom, SeokJin</u> (Seal)

Abstract

A study on Validity of Mongolian Civil Service Entrance Examination for Job Performance

Amartuvshin Amarmend

Public Administration Major

The Graduate School of Public Administration

Master of Public Administration

Seoul National University

The validity of Mongolian civil service entrance examination for job performance has been tested using data from Civil Service Council of Mongolia. Data sources are entrance examination database and Human Resource Management System (HRMS) which includes general human resource data of all Mongolian government officials. Multiple linear regression has been chosen to estimate parameters in regression model. Assumptions of multiple linear regression have been checked and it is considered Ordinary Least Squares (OLS) method is acceptable for this research. Main dependent variable was a performance evaluation score of the public servants and main independent variable was their score that they get on civil service entrance examination. Age, education, gender, experience, ranking and degree of subjects' position, type and size of the organization that

they are working for, IT and language literacy of subject, and awards they received is included in the regression model as a control variable.

The results of this study showed that Mongolian civil service entrance examination is valid for job performance with parameter estimated β =0.059 at 0.01 significant level, (p = 0.0006). Among the components of the entrance examination, document examination score showed strongest predictor (β =0.284, p =0.0001) and competence examination was estimated second strongest predictor (β =0.18, p =0.0171) for job performance. Multiple choice test examination score did not show significant relationship with job performance level (β =0.031, p = 0.1196). Individual interview also did not show significant relationship with job performance (β =0.036, p = 0.4243).

The results were consistent with findings from focus group interview results. Implications to improve entrance examination have given in conclusion part.

Keywords: Validity, Entrance examination, Job performance, Job performance predictor, Multiple regression, Mongolia.

Student ID: 2012-24069

Table of Contents

Chapter I. INTRODUCTION

1.1. Study Background

Mongolia has changed from socialist country to market economy after 1989. We changed from a party centered country to state centered country. Since then we started to build a new country. In every sector of the society, in every branch of the law there has needed a much to be done in order to create state centered free market economy. One of these goals was creating a new merit based civil service in my country. In general for the developing countries efficiency of their government, good governance, and corruption in the government is very important to decide their development trajectories. In order to build a new merit based civil service system Mongolia has created a new organization, Administrative Service Council, in 1995.

When creating merit based civil service system Mongolia has faced many difficulties but merit based selection system was one of the most successfully created system. Recruitment and selection is the first stage of human resource management process. Competitive selection process which is solely based on the performance on the examination results of applicants in civil service entrance examination would be a base for later stages of human resource management. Amongst the principles of merit based civil service system merit based recruitment principle is one of most important principle.

Moreover, the examination result should be related to job performance. Applicants who have higher score on the test should show

higher performance after they recruited. That is the whole point of conducting competitive test to select best and brightest. This logic is also included in the related laws and legislations. For example: In article 17.3 of Mongolian Civil Service Law, CSC conducts examination and nominates a candidate who meets best for the special and general requirements for respective position.

From practical point of view, this topic was easier to approach because there were relatively adequate data about the test scores of the applicants and performance evaluation scores of the current workers compared to the data about other aspects of human resource management in Mongolia. I worked for three years in Civil Service Council of Mongolia which conducts Civil Service Examination and collects data for current employees. This research was possible because I had an access to those data which is very hard for general public or researchers to access. Because of the lack of data there was no empirical studies conducted in this field but some research have been done in theoretical perspective.

1.2. Purpose of Research

Generally it is still hard to say that there is a sufficient study about Mongolian civil service system after transition to market economy. To contribute to the research field on contemporary public administration and to fill the gap that caused from the lack of data in general and lack of access to data existing for general public is the purposes of this research.

This study will focus on the civil service examination and try to answer the questions:

-Does civil service entrance examination valid to predict future performance of government officials? If it is valid how strong is the prediction power of the examination? Among the components of civil service examination which component is most relevant to performance on the job?

Moreover, there is a public criticism that civil service entrance examination only focuses on the knowledge about law and regulation and needs to be diversified for the various types of job in government. With this study we can check whether the criticism is acceptable or not. We could discuss more deeply based on the empirical findings, whether we diversify the entrance examination specifically designed to different types of jobs rather just argue based on hunt or well established logic.

1.3. Methodology

This research tried to calculate the predictive validity of civil service examination for job performance, using data for civil service examination and civil service human resource database for 2009 to 2012 which includes data for performance evaluation of government officials. Both of these databases are provided from Civil Service Council of Mongolia. A quantitative methodology was followed. Multiple linear regression is used for research method. Ordinary Least Squares (OLS) method has been chosen to estimate parameters in regression model. Main dependent variable is a performance evaluation score of the public servants. Main independent variable is their score that they get on civil service entrance examination.

Chapter II. THEORETICAL BACKGROUND

2.1. Merit based Civil Service system

2.1.1. Merit based system

When a country's civil service is based on merit system, then civil service can contribute to its country's development more and considered as more efficient. There are many different definition of merit system. According to O.Glenn Stahl definition "In broadest sense a merit system in modern government means a personnel system in which comparative merit or achievement governs each individual's selection and progress in the service and in which the conditions and rewards of performance contribute to competency and continuity of the service". Merit system is based on the ideal - the belief that a competent, committed workforce of career civil servants is essential to the professional conduct of the public's business. It has created as a result of increased pressures for rational and transparent government and increased demands for more effective delivery of public services. In US Pendleton Act(1883) marked fundamental shift from patronage to merit system. In simple words the recruitment, promotion and all other aspects of human resource management should be exclusively based on their job performance and ability, rather than their political connections. It is the opposite of the spoils system which means a practice where a party distributes government jobs to people as a reward for their work in the election campaign, vote or just

for the personal gain. This system has been popular in United States, UK, Australia and New Zealand in their early stage of modern development. It grants political responsiveness.

Merit system is based on 9 principles:

(1) Recruitment should be from qualified individuals from appropriate sources in an endeavor to achieve a work force from all segments of society, and selection and advancement should be determined solely on the basis of relative ability, knowledge and skills, after fair and open competition which assures that all receive equal opportunity.

(2) All employees and applicants for employment should receive fair and equitable treatment in all aspects of personnel management without regard to political affiliation, race, color, religion, national origin, sex, marital status, age, or handicapping condition, and with proper regard for their privacy and constitutional rights.

(3) Equal pay should be provided for work of equal value, with appropriate consideration of both national and local rates paid by employers in the private sector, and appropriate incentives and recognition should be provided for excellence in performance.

(4) All employees should maintain high standards of integrity, conduct, and concern for the public interest.

(5) The Federal work force should be used efficiently and effectively.

(6) Employees should be retained on the basis of adequacy of their performance, inadequate performance should be corrected, and employees should be separated who cannot or will not improve their performance to meet required standards.

(7) Employees should be provided effective education and training in cases in which such education and training would result in better organizational and individual performance.

(8) Employees should be:

    a.   Protected against arbitrary action, personal favoritism, or coercion for partisan political purposes.

    b.   Prohibited from using their official authority or influence for interfering with or affecting the result of an election or a nomination for election.

(9) Employees should be protected against reprisal for the lawful disclosure of information which the employees reasonably believe as evidence:

    a.   A violation of law, rule or regulation.

    b.   Mismanagement, a gross waste of funds, an abuse of authority, or a substantial and specific danger of public health or safety.

We can see building a merit based system was not an easy task for many countries. In the beginning of 1880s Britain, USA and Australian public service were plagued by problems of corruption and inefficiency. President

George Washington prized his partners who fought in war of independence with important positions in civil service and President Abraham Lincoln gave a priority for unity in his appointments during civil war rather than an ability and performance. But after a few decades of practice and it widening to the all posts of civil service were filled through political patronage, government was unable to perform its duties in the face of development where corruption, inefficiency and working for private gain is common. In Britain famous Northcote-Trevelyan Report in 1853, which provided a blueprint for the transformation from a patronage system. The report noted that government:

> "…could not be carried on without the aid of an efficient body of permanent officers … possessing sufficient independence, character, ability, and experience to be able to advise, assist, and to some extent, influence, those who are from time to time set over them."

Public services were reformed to foster a politically neutral career service. In USA Pendleton Act, an act to regulate and improve the civil service, was passed in 1883.

The major elements of the merit system today is standardized conditions of employment, administration by an independent agency, merit-based recruitment through competitive examination, merit-based promotion and redundancy not affected by political affection.

2.1.2. Merit Based Recruitment System

One crucial part of the merit system is competitive selection process which is solely based on the performance on the test of applicants which is mentioned above in the first principle of the nine merit principles:

"Recruitment should be from qualified individuals from appropriate sources in an endeavor to achieve a work force from all segments of society, and selection and advancement should be determined solely on the basis of relative ability, knowledge and skills, after fair and open competition which assures that all receive equal opportunity."

Selection process is a first stage of career for the employee in public service. If we cannot set appropriate selection process later stages of merit system would be less meaningful. That is the one reason I want to focus on selection process especially civil service entrance examination in Mongolia.

2.2. Literature Review

2.2.1. Merit based Civil Service System in Mongolia

In a new constitution(1992) after the end of communist era, article 1 section 2 stipulated "The primary principles of civil service shall be ensuring of democracy, justice, equality, national unity and respect for law." Also article 46 section 2 stipulated as "State employees shall be Mongolian citizens. They shall strictly abide by the Constitution and other laws and work for the benefit of the people and in the interest of the State." and section 3 as "The working

conditions and social guarantees of State employees shall be determined by law."

Based on these regulations in constitution New Civil Service Law has been adopted in 1994 and long and uneasy process of building merit based civil service started. Upon this Civil Service Law the independent agency, Administrative Service Council of Mongolia, responsible for "planning the implementation of the Civil Service Reform Strategy, provide with the policy guidance the state human resources, coordinate and organize implementation thereof" has created. The parliament has readopted Civil Service Law in 2002 and made Administrative Service Council into Civil Service Council (CSC) giving more mandate, wider range of public organizations under their jurisdiction. Also in 2008 amended Civil Service Law aiming to reassure professional merit based civil service, to ensure civil service to be free from unlawful political influence and to assure independence of Civil Service Council. Article 34.1 of Civil Service Law provides "Civil Service Council (hereinafter "Council") shall be the central civil service body." and article 34.2 puts down "The Council shall be an independent body accountable to the Parliament."

The article 4 of the Civil Service Law provides that: "Unless otherwise provided in law, the following principles shall be observed in civil service:

4.2.1.to manage and to be subordinated;

4.2.2.transparency;

4.2.3. serving the people;

4.2.4. equal opportunity for citizens to enter civil service on the conditions and rules defined by law;

4.2.5. professionalism and stability of civil service;

4.2.6. the state ensuring conditions and guarantees to civil servants for the full exercise of their official powers;

4.2.7. the state being liable for the damage caused by civil servants in the course of exercising their powers afforded by law."

Since the declaration of People's Republic, Mongolia has been under strong Soviet influence. We have been in a socialist regime for almost 65 years. Under socialist regime those who are loyal to the communist party and who has knowledge, skill and ability and shows good performance had a chance to get promoted. The system was strongly controlled by party interests although inside the party there have been well established human resource management policies in practice. In other words as long as the ideology does not matter, and person has no conflict with party we had a quite good system of management in public sector. However, after the shift to market economy that party centered merit based human resource management system has lost.

Civil service law 1994 tried to implement merit system in Mongolia. Although, it is hard to observe that we successfully created merit based civil service in Mongolia. Especially political influence to civil service on hiring, promotion and dismissal is critical. There are a ways to circumambulate the

merit system and put their own people into civil service. Theoretically employment security enabled a professional and impartial public service to provide frank advice without fear of dismissal. A non-partisan civil service is able to serve any government and do not need to be dismissed upon a change of government. They are professional who must aid the elected officials to implement policy but not political actors, using civil post for their private or party interest.

After 1990s and globalization, many new viewpoint about merit system is introduced in many countries, for example concerns regarding efficiency, managerial control and political responsiveness, and many countries have changed their merit system and even some disposed of their independent agency for merit system. However, it does not mean that developing countries like Mongolia should aim for more political responsive, more patronage oriented civil service. Those developed countries already had well established merit based civil service before recent years reform and they basically loosened strict merit based system that they initially built in order to improve flexibility in government human resource management. But developing countries does not have such precondition they are on the way to build a good civil service based on scientific management and merit system. We cannot imitate developed countries directly, bypassing a certain development stages.

2.2.2. Entrance examination

Human Resource Management (HRM) is fundamentally about matching human resources to the strategic and operational needs of an organization and ensuring full utilization of those resources. It is concerned not only with obtaining and keeping the right number and quality of staff required but also with selecting and promoting people who "fit" the culture and strategic requirements of the organization (Bello-Imam, Oshieonebo and Ojeifo, 2007).

Today, the use of tests for employment screening and placement is widely used. Private and public almost all the organizations use some form of test. Psychological tests are widely used in industry as an aid in the selection and placement of employees. Employees are vital asset in organizations. That is why the process of hiring needs to be paid serious attention as a means to get those vital assets.

In order to select a right person for the job, individual differences in terms of abilities and skills need to be adequately and accurately measured for comparison. This is done through a device called selection test or employment test (examination). A test is any device used to separate qualified from unqualified applicants for selection or promotion. Written examination and interviews, medical examinations, drug tests, background investigations, and physical requirement can be included in tests.

There are various types of test such as: Cognitive ability tests, Mechanical ability tests, Motor and sensory ability tests, Aptitude Tests, Achievement Test, Job skills and knowledge tests, Personality tests, and Miscellaneous tests.

12

A. Concept of Validity

An employment test is a sample of a person's job related behavior. Its main aim is to measure the qualities and characteristics which lead to successful performance on the job. A good employment test should have the following characteristics:

1. Reliability

In personnel measurement, the term reliability means the dependability or consistency of a measure. To explain reliability, it is easiest to begin by considering an unreliable measure to be one that produces one set of results on one occasion and a different set of results on another occasion, with nothing to explain the differences. (Cronbach, 1970).

On the other hand, when a reliable measure is applied repeatedly to a single person, it will yield the same number each time it is applied.

The purpose of having high reliability in personnel measures is to be able to arrive at sound judgments about the people being measured. With high reliability, we can be more confident that a particular measure is giving a dependable picture of an individual's true score.

2. Validity

The term validity refers to how well a measure actually assesses the attribute it is being used to asses. Personnel selection is the degree to which that measure actually assesses suitability for the job in question.

The validity of a measure is not absolute but is relative to the use to which it is put. It is valid as long as it successfully measures the attribute that the user intends it to measure. In other words, a measure that is valid for one purpose

may not be valid for another. To be valid the test must separate more-qualified from less-qualified applicants on the basis of job-related competencies or performance. For example, a test process is valid if those who do well on the test do well on the job, and those who do not do well in the test do not do well on the job.

		Job Performance	
		High	Low
Test Score	High	**True Positive**	False Positive
	Low	False Negative	**True Negative**

This table shows in matrix form the possible relationships between test scores and job performance. The more valid a test, the more **true positives** and **true negatives** it will produce. In other words, those who do well on the test in fact will do well on the job, and those who do poorly on the test will do poorly on the job. The less valid test, the more likely it will produce **false positives** and **false negatives**. A false positive is a person who does well on the test but does not do well on the job. A false negative occurs where a person fails to do well on the test but does well on the job. The test falsely predicted the applicant's performance.

Types of validity

- Predictive validity – the extent to which the test correctly predicts future behavior. To establish predictive validity it is necessary to conduct extensive research over a period of time. It is also necessary

14

to have accurate measure of performance so that the prediction can be compared with actual behavior.

- Concurrent validity – the extent to which the test score differentiates individuals in relation to a criterion or standard of performance external to the test. This means comparing the test scores of high and low performances as indicated by the criteria and establishing the degree to which the test indicates who should fit into the high or low performance groups.

- Content validity – the extent to which the test is clearly related to the characteristics of the job or role for which is being used as a measuring instrument.

- Face validity – the extent to which it is felt that the test 'looks' right, is measuring what it is supposed to measure.

- Construct validity – the extent to which the test measures a particular construct or characteristic.

2.2.3. Job Performance

There are various approaches to define job performance such as micro definitions that focus on specific individual behaviors and macro definitions that focus on productivity or effectiveness. However, generally researchers agree that job performance can be defined on a micro level as "actions and behaviors of an employee that contribute to the goals of the organization" (Campbell, 1990; Murphy, 1989).

Rotundo (2000) explained that although researchers provide their own conceptualization of job performance, a typical definition of job performance focuses on behaviors or actions of individuals, not results or outcomes of these actions and behaviors. Murphy (1989) states that job performance should be defined in terms of behaviors rather than results. He explained that result-based measures are not always functional to the organization, as employees may try to maximize results at the expense of other things. For example, employee may break the work rules and achieve high result. Moreover, when the results of employee related to forces outside of the organization or his/her influence, as the outer environment changes consistent effort or behavior does not lead to same results.

Job performance can be described by three broad categories of behaviors, which are subsequently labeled task performance, organizational citizenship, and counterproductive performance (Rotundo, 2000).First category means behaviors associated with the production of a good or the provision of a service. Some researchers propose that employees do not spend all of their time at work performing activities that are strictly related to task performance (e.g., Borman & Motowidlo, 1993; Katz & Kahn, 1978; Organ, 1988).Employees help coworkers or volunteer to engage in activities that benefit the organization. These types of behaviors are called organizational citizenship behavior or contextual performance. The third category or domain is counterproductive behavior. It represents negative behaviors such as tardiness, theft, sabotage, absenteeism, etc. that can harm the well-being of the organization or coworkers (Crino, 1994; Giacalone & Greenberg, 1997).

A. Measurement of Performance

The most common way to measure job performance is a supervisor or manager's rating of an employee's job performance. Also production records, sales records, and other measures were also used. Examples of studies measured job performance as a supervisor rating (or peer rating) include (Schmidt, Hunter, and Outerbridge 1986), (Borman et al.,1991), Hunt (1996) etc. On the other hand there are studies that measured performance as more complex set of terms. Kahya (2007) used 7 criteria to measure task performance. Criteria included: Job knowledge, Overcoming obstacles to complete a task, Problem solving, Operating equipment, using tools, Working safely, Concentrating to the duties, and Protecting the resources.

Kahya (2007) also discussed 16 criteria in three groups to measure contextual performance based on previous studies.

For Interpersonal Citizenship:

- Assisting co-workers with personal matters,

- Cooperating with others to solve problems,

- Engaging responsibly in meetings and group activities,

For Organizational Citizenship:

- Treating the supervisor with respect,

- Absenteeism,

- Working systematically,

- Following organization rules and proper procedures,

- Participating responsibility in the organization,

- Completing a task on time,

For Job Dedication:

- Attention to important details for job,

- Quality,

- Productivity,

- Creativity to solve a work problem,

- Engaging in self-development to improve own effectiveness,

- Generating new ideas to make things (tasks) better (innovation),

- Planning and organizing work.

Supervisor ratings of employee performance are considered as subjective measures, while objective measures are considered to be "direct measures of countable behaviors or outcomes" (Bommer et al., 1995).

There are few types of performance measurement systems in practice. In 360-degree performance measurement system performance appraisals of the target employee are provided by supervisors, peers, subordinates and employee himself. In some organizations the feedback from supervisors or peers are used. Although some research indicates that self-ratings of job performance are congruent with supervisory ratings of job performance (e.g.,

Farh, Werbel, & Bedeian, 1988) other research indicates the opposite (McEnery & Blanchard, 1999).

B. Determinants of Supervisory Ratings on Job Performance

Job performance is one of most studied dependent variable in the fields of human resource management as well as industrial and organizational psychology. Various kinds of variables have studied such as sex(Lee and Alvares, 1977), rater's opportunity to observe rate's job performance, rater's supervisory experience (Moser et al, 1999, Antonioni and Park,2001), interpersonal affect(Antonioni and Park,2001), rating format(Yun et al., 2005), organizational politics (Witt et al., 2002), and workplace deviant behaviors(Dunlop and Lee, 2004) as determinants of supervisory ratings of job performance. These factors seem to affect the supervisory rating process for job performance. However, they do not directly define the job performance of a worker.

Lee and Alvares (1977) examined the effect of subordinate and supervisor sex on the description and evaluation of supervisory behavior. They conducted laboratory simulation with participants of sixty four male and sixty four female subordinates and four male and four female supervisors who trained to exhibit specific supervisory behaviors. They found no differences in the descriptions and evaluations as a function of the sex of the supervisor. The sex of the supervisor was not an important determinant of supervisory descriptions and evaluations.

Moser et al (1999) studied the relationship between validity of assessment center and opportunity to observe. The uncorrected validity of assessment center was found 0.37 but opportunity to observe calculated as an amount of time the rater was ratees' supervisor found a moderator. Assessment center validities were 0.09 when opportunity to observe is equal or less than 2 years, while validities were 0.50 when opportunity to observe is greater than 2 years.

Antonioni and Park (2001) investigated the strictness difference of the rating among three types of evaluation: downward, upward, and peer. Also they checked the effect of observation time as rater. They found the influence of rater effect on the strictness of ratings was significantly lower in upward and peer feedback than in downward feedback and that the influence increased as raters' observation time increased.

Yun et al., (2005)investigated the effects of rater personality (Conscientiousness and Agreeableness),rating format (graphic rating scale vs. behavioral checklist), and the rating social context (face-to-face feedback vs. no face-to-face feedback) on rating elevation of performance ratings. They found that raters with high agreeableness showed more elevated ratings than those with low agreeableness when they expected to have the face-to-face feedback meeting. According to their result rating format moderated the relationship between agreeableness and rating elevation.

Witt et al. (2002) explored the process of evaluating contextual performance in the context of apolitically charged atmosphere. They studied matched sample of 540 supervisors and subordinates employed in the private

sector. They found that the interaction of politics and the personality dimension of agreeableness explained a significant incremental amount of variance in the interpersonal facilitation facet of contextual performance.

Finally, as an example of possible relevant determinant of general job performance study of Dunlop and Lee (2004) can be introduced. Dunlop and Lee (2004) found that the presence of workplace deviant behavioral employees among business units impinges upon the performance of the business unit as a whole. However, the organizational citizenship behavior, opposite of deviant behavior, had comparatively little effect for performance as a whole.

2.2.4. Validity Research Findings and Meta-analyses

Determinants discussed in previous section are not related to individual job performance itself but they affect workers job performance as a supervisor ranking indirectly. Depending on the determinants discussed above workers with same performance may get different ratings. Those are more of like third variable to performance ratings as indicator of individual job performance. However, in this section we focus on direct determinants or predictors of individual job performance.

There have been literally thousands of studies of the validity of selection tests. There have been also studies for reviews of validation research. (Hunter & Hirsh, 1987; Hunter & Hunter, 1984; Schmidt, Ones, &Hunter, 1992; Schmidt and Hunter, 1998)

Schmidt and Hunter (1998) offered comprehensive summary of practical and theoretical implications of 85 years of research in personnel selection. Research studies assessing the ability of personnel assessment methods to predict future job performance and future learning(e.g., in training programs) have been conducted since the first decade of the 20th century (Schmidt and Hunter, 1998). Since then thousands of studies have conducted in the field of validity study. However, early studies showed different results even though they are conducted on a same subject. It was discovered that most of the differences across studies were due to statistical and measurement artifacts such as sampling error due to small sample and not to real differences in the jobs (Schmidt and Hunter, 1977; Schmidt, Hunter, Pearlman, and Shane, 1979).

As the validity studies accumulate focus of studies shifted to the generalization of validity. In other words they tested whether predictor valid for job performance in particular type of job is valid for other types of jobs. Number of predictors has been found general to jobs. Moreover, new techniques to combine results of validity study in systematical way invented. Glass (1976) has recommended a method called meta-analysis to that could combine validity estimates across studies. The simplest form of meta-analysis is to average correlations across studies. More advanced formulas of meta-analysis have also developed by Hunter, Schmidt and Jackson (1982).

Studies based on meta-analysis method provided more accurate estimates of the average operational validity and showed that the level of real

variability of validities was usually quite small and might in fact be zero (Schmidt, 1992).Ghiselli (1973) summarized thousands of studies done over the first 60years of validity study. He checked mean validity of three predictors such as general cognitive ability, general perceptual ability and general psychomotor ability among different types of job families. Dunnette (1972) has reviewed 883 validity studies and counted correlations of predictors with job performance. He used predictors such as cognitive ability, perceptual ability, psychomotor ability, biographical inventories, interviews, education, job knowledge and job tryouts. Reillyand Chao (1982)and Vineberg and Joyner(1982) also reviewed 103 and 246 studies, respectively.

Schmidt and Hunter (1998) summarized main meta-analyses conducted and compared variability of 19 selection methods for predicting performance on the job. They have also calculated validity for these method's for job related learning. They also calculated the validity of paired combinations of general mental ability (GMA) and the 18 other selection procedures. Their main findings are as follows:

Table 1.

Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores Combined With a Second Predictor Using (Standardized) Multiple Regression

Personnel measures	Validity (*r*)	Multiple *R*	Gain in validity from adding supplement	% increase in validity	Standardized regression weights	
					GMA	Supplement
GMA tests[a]	.51					
Work sample tests[b]	.54	.63	.12	24%	.36	.41
Integrity tests[c]	.41	.65	.14	27%	.51	.41
Conscientiousness tests[d]	.31	.60	.09	18%	.51	.31
Employment interviews (structured)[e]	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured)[f]	.38	.55	.04	8%	.43	.22
Job knowledge tests[g]	.48	.58	.07	14%	.36	.31
Job tryout procedure[h]	.44	.58	.07	14%	.40	.20
Peer ratings[i]	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method[j]	.45	.58	.07	14%	.39	.31
Reference checks[k]	.26	.57	.06	12%	.51	.26
Job experience (years)[l]	.18	.54	.03	6%	.51	.18
Biographical data measures[m]	.35	.52	.01	2%	.45	.13
Assessment centers[n]	.37	.53	.02	4%	.43	.15
T & E point method[o]	.11	.52	.01	2%	.39	.29
Years of education[p]	.10	.52	.01	2%	.51	.10
Interests[q]	.10	.52	.01	2%	.51	.10
Graphology[r]	.02	.51	.00	0%	.51	.02
Age[s]	−.01	.51	.00	0%	.51	−.01

Among this predictors they have put most emphasis on GMA because amongst the processes which could be used for all levels of jobs, it has highest validity with lowest cost, in many of meta-analyses they have summarized GMA was the strongest predictor in most of them. GMA has also been shown to be the best predictor of job related training and it also has stronger theoretical foundation. Calculating combined validity of other predictors with GMA is another originality of their work. They also found that GMA has a highest validity for job related learning .56.

Employment interviews can be either structured or unstructured. Unstructured interviews have no fixed format or set of questions to be answered. As a result, the same interviewer often asks different applicants different questions. Structured interviews have fixed format and all questions are prepared beforehand and asked in the same order to each of interviewee. As shown in previous table from Schmidt and Hunter (1998), the average validity of the structured interview is .51, and validity for the unstructured interview is .38. But they also found that these two types of interview had same validity for job related learning .35.

Job experience has been found with only .18 validity for job performance in average. They explained this finding is a result of the variety of performance from less than 6 months to more than 30 years in their data. However, Schmidt, Hunter, and Outerbridge (1986) found that when experience on the job does not exceed 5 years validity for job experience was .33. Job experience exceeding 5 years did not result increase in job performance.

Biodata measures are generally comprised of questions related to applicants' personal backgrounds and life experiences. Lautenschlager and Shaffer (1987) defined biodata as "...factual kinds of questions about life and work experiences, as well as to items involving opinions, values, beliefs, and attitudes that reflect a historical perspective." The basis of biodata's predictive abilities is the axiom that past behavior is the best predictor of future behavior. However, biodata measures are technically difficult and time

consuming to construct and considerable statistical sophistication is required to develop them. From the findings of Schmidt and Hunter (1998) above, biodata has validity of .35 for job performance. Also they found that biodata has validity for job related training .30.

Education has been found a weak predictor for job performance with .10 validity. They explained this value as "It is important to note that this finding does not imply that education is irrelevant to occupational success; education is clearly an important determinant of the level of job the individual can obtain. What this finding shows is that among those who apply to get a particular job years of education does not predict future performance on that job very well." However, education has also been found .20 validity for job related learning.

Age has been found not valid for predicting job performance. Although one may suppose age to be related with job performance but here is shown that age is not related to it. It has also been found that age is not related to job related learning as well.

In reading the results shown in previous table we should remember that these results are found with meta-analyses. That means these are the average of findings of previous studies. Individual findings of each predictor's validity have different range. However, these results from meta-analyses give us some approximate estimation of real validity.

Maybe most important finding from their work is that they explained the underlying psychological process why these predictors are related to better performance. They suggested that major reason that GMA was found to be good predictor is the impact of mental ability to the acquisition of job knowledge. More intelligent people have higher job performance because that they acquire job knowledge more rapidly and acquire more of it. Thus, mental ability has its most important effect on job performance indirectly, through job knowledge. There is also a direct effect of mental ability on job performance independent of job knowledge, but smaller. For nonsupervisory jobs, this direct effect is only about 20% as large as the indirect effect; for supervisory jobs, it is about 50% as large (Borman et al., 1991; Schmidt, Hunter and Outerbridge, 1986).

Finally, we should pay attention to the combination results of selection procedures with GMA since in most practical cases more than one kind of selection procedure is used. Third and fourth column of their finding shows these results. There is a correlation between predictors. Schmidt and Hunter (1998) focused on the correlation of other predictors with GMA. Some predictors such as work sample test, integrity test, and structured interview showed relatively smaller correlation with GMA so the bigger gain in validity when they combined with GMA has been achieved. On the other hand some predictors such as biographical data measure and assessment centers showed relatively high correlation with GMA so they gain from the combination with GMA was very low.

There is also some research that lower-ability people can actually be more effective performers under certain conditions such as time pressure, social evaluation, and unpredictable task change. Gimmig et al.(2006) suggested "individuals who are most likely to fail under performance pressure are those with high working memory capacity". They also offered that failing under pressure is not confined to tasks involving acquired skills and knowledge. Beilock and DeCaro (2007) conducted experiments of mathematical problem solving under pressure and found under high-pressure conditions, higher working memory individuals used simpler (and less efficacious) problem-solving strategies, and their performance accuracy suffered. Lang and Bliese (2009) researched relationship between general mental ability and adaptive performance by applying a discontinuous growth modeling framework to a study of unforeseen change on a complex decision making task. They found that GMA was negatively related to transition adaptation and found no evidence for a relationship between GMA and reacquisition adaptation meaning loss of performance was bigger for high GMA people and also they were not as adaptive as expected. DeCaro, Thomas, and Beilock (2008) showed that for information-integration categories which believed to be driven by procedural learning processes that operate largely outside of conscious control, lower-ability people learned faster than one's with high working memory.

Chapter III. OVERVIEW OF MONGOLIAN SYSTEMS

3.1. Stages to fill vacancy in government positions

Stages to fill vacancy have defined in Civil Service Law article 17 and CSC Decree concerning Civil Service competence examination (2009)[1]:

1. When there is a vacancy in government organization the organization first should seek for appropriate personnel to fill the position within their or other similar organization.

2. If there is nobody available then the organization should request CSC (Civil Service Council) to nominate possible candidate from the reserve list of previous examinations.

3. If there is no one appropriate for the position according to previous 2 steps then the organization send a request to recruit personnel through examination to CSC. CSC conducts examination and nominates a candidate who meets best for the special and general requirements for respective position[2] as stipulated in CSL article 17.3.

This procedure means that the civil service entrance examination could be conducted as many times as there has been request from the government organizations to the CSC. If there are many requests to fill the

---

[1] The Decree has revised in May 2013, but this study has covered the examination data that is conducted in previous Decree concerning Civil Service entrance examination of 2009. Moreover, official nomenclature of Civil Service entrance examination is "Civil Service competence examination".

[2] In practice this article is interpreted as a candidate who marked highest score on entrance examination.

vacancy greater number of examinations will be conducted and if there is no request, the examination will not be conducted. Administrative and special civil service positions are required to take civil service examination.

3.1.1. Components of Civil Service Entrance Examination

Components of the examination have set by CSC Decree concerning Civil Service competence examination (2009):

a. Multiple choice test examination. This examination evaluates the knowledge of laws and regulation about government and its policy. The multiple choice test part of the examination consists of 60 questions and each correct answer to questions is given 1 point. Duration of this part is 60 minutes. Those who got 35 or more points on this part of examination may proceed to next parts of the examination.

b. Document examination. In document examination examiner evaluates the documents provided by candidates showing their fit to the general requirements of the position regarding education, profession, competence and experience. If Grade Point Average (GPA) score of the candidate's diploma of bachelor or equivalent is not less than 2.0 and candidate meets other 3 requirements of position candidate is given "Pass" for this part of examination. If one of these requiments is not met candidate is given "Fail" for this part of examination. For candidates passed document examination points are given in relevant to candidate's education. For bachelor diploma or equivalent GPA score is given (2.0-4.0 points). For higher degrees of education 2 to 6 points is

given according to following table. If the candidate possesses more than 2 degrees only highest degree is scored. Maximus point for document examination is 10 points.

Table 2.

Education level	Bachelor or equivalent	Master		Ph.D		Sc.D
		Domestic	Foreign	Domestic	Foreign	-
Points	GPA 2.0-4.0	2	3	4	5	6

c. Individual interview examination. 5 questions asked concerning relevant organization and position's mission, goal, responsibility and required knowledge, attitude and behaviour for the position. Each question is scored 1 to 3 points. Maximum 15 points.

d. Competence examination. In regards of general requirements for the position, candidate's competence of official documentation standard and others is evaluated by 10 points and knowledge for widely used computer programs is evaluated by 5 points. Maximum 15 points.

Total possible maximum point for these 4 components is equal to 100 points.

3.2. Performance evaluation in government organizations

In 2002 Public Service Management and Finance Law (PSMFL) has been adopted. This law aimed to introduce New Public Management principles into government. According to PSMFL every public organization has general manager and portfolio minister. General manager is a manager who oversight

daily management of organization. General manager is responsible before portfolio minister. "Portfolio Minister" means an official who within his/her delegated authorities determines the need for outputs to be delivered by budgetary or other bodies, and who is entitled to conclude output purchase agreement. Portfolio minister himself is responsible before the citizens. That's the main principle of PSMFL regarding performance evaluation. In 2002 relevant changes have been made to CSL too.

According to CSL article 19 CSC has a mandat to establish establish the procedure for evaluation of the performance and qualification level of the Core civil servants, unless otherwise provided for in law.CSC has established "Decree concerning the compact of result of contruct for government officials and evaluation of performance level and competency level of government officials" in 2007[3]. This decree regulates performance evaluation for all the administrative and special civil servants. All administrative civil servants signs a performance contract with the managers in their organization. Tasks to be done by the personnel is defined in consistent with the long term plan of the organization whom working for and organization and unit's operation plan for following year, and job description of relevant position. General manager of the organization should consult with the official drafting result of contruct and provides proper advice. In addition to the tasks set forth, criteria for the successful completion such as quantity, quality, timeline and workingdays are set as well.

---

[3]This decree has revised in October 2012 which came to in effect from January 2013. All the performance evaluation data included in this thesis has been regulated by 2007 Decree since forth focus on 2007 Decree.

Performance evaluation of a official consists of 3 parts.

1) General tasks. Tasks related to the official's responsibility in accordance with the organization's strategic and operation goals are listed as general tasks. Other tasks that is not included in special tasks are also included. Fulfilment of each task is evaluated up to 40 points and average of these tasks is calculated.

2) Special tasks. The tasks that contributes to implement the governmen policies, to help to fulfill general manager's special tasks, tasks directly connected to organization's primary goal and other tasks considered as most important by general manager are included. Fulfilment of each task is evaluated up to 30 points and average of these tasks is calculated.

3) Competence improvement related tasks. Based on the knowledge and competencies included in job description of employee, tasks to improve individual competence and knowledge is included in consultation with general manager. Comply the rules and standards of government and ethic code might be chosen as a goal. Fulfilment of each task is evaluated up to 30 points and average of these tasks is calculated.   All these tasks set in the begining of the year is evaluated in the end of the year. Completion of the performance contract should be reported quarterly to the managers but evaluated yearly basis.

Average of above 3 types of evaluation is calculated and sum of the average is the performance evaluation point for the employee for that year. 0-59 is market as "E", 60-69 is marked as "D", 70-79 is marked as "C", 80-89 is marked as "B", and 90-100 is marked as "A". However these evaluations

should be given by the supervisor according to the Decree, in practice evaluation points is first done by self-evaluation as a task fulfillment report to the supervisor. Then supervisor checks the report and make corrections when necessary. In some cases when supervisor have to evaluate number of subordinates, they could not go thorough about the report. So performance appraisal in Mongolian government is not pure supervisor rating, kind of mixture in practice. Depending on the mark of the evaluation direct evaluator of the contract can make recommendation to improve salary, to train, to give bonus, to demote, to use disciplinary action, and to dismiss for the higher level manager. However, common view among government employees is that performance evaluation is just a symbolic.

Chapter IV. RESEARCH DESIGN

As selection is one crucial precondition of the merit system and performance of the government officials and availability of data was relatively abundant to concerning civil service entrance examination this thesis especially focus on scores of the examinee's on civil service entrance examination as a predictor for their job performance after they appointed to the service.

4.1. Research Hypothesis

Civil Service Entrance Examination is conducted in order to fill the vacancy in government position with a candidate who meets best for the special and general requirements for respective position as stipulated in CSL article 17.3. In other words the candidate who marked highest result on examination gets appointed to the position. That is a commonly accepted practice and it is the result of the expectation that candidate who performs well on examination will likely to perform well on the real job too.

That's why main hypothesis of this study is:

Hypothesis 1: The higher the examination score the higher will be performance evaluation score. Performance evaluation score and entrance examination score would have positive relationship.

Additionally hypothesis 2 will be checked:

Hypothesis 2: Amongst the components of civil service entrance examination the score on multiple choice examination would be most strongly related to job performance compared to other components of entrance examination. This component is taken most seriously by both officials and candidates and its weight in the examination score is highest at 60 out of 100.

4.2. Methods for analysis

A quantitative methodology was followed. Multiple linear regression method has been chosen as a research method. Ordinary Least Square methodology is used. Main dependent variable is a performance evaluation score of the public servants. Main independent variable is their score they get on civil service entrance examination. Possible factors that affect examination score or job performance evaluation ratings have been checked and included as a third factor when data is available. Regression assumptions have checked it is considered assumptions of OLS method are met. Cautions have been added when necessary. For regression calculations SAS 9.3 TS Level 1M0 version of the SAS program is used.

4.3. Data used in Research

4.3.1. Data source for Independent variable

Data for candidates who took civil service entrance examination and registered for the reserve list of the position is collected for year 2009, 2010, 2011 and 2012. Each time when there is an examination conducted the list of

people for the reserve has been created and saved in a MS-Excel file containing the detailed information related to the examination such as information about the position, candidate's personal information (name, age, gender, address, citizenship ID, contact info), candidate's education, major, qualifications, results of the examination.

4.3.2. Data source for Dependent variable

Data for performance evaluation ratings have obtained from human resource management database which is collected through HRMS program. According to CSL article 37 and Decree of CSC 17 of 2004 every government organizations should send reports of their human resource management data to CSC in the end of each year. The HRMS database is in Mysql database format. Data included in HRMS are: information about organization; information about workers employment status; detailed personal information on worker; information about family members; information about education; information about training courses attended; self-evaluation of general, IT, and language skills; work experience; detailed information about the position; payment; performance evaluation and other human resource management related information such as whether the worker made some initiatives, received government awards, etc. The data in the database is updated and cumulated in each year. For this study HRMS database dated 2012.12.31 has been used.

4.3.3. Data Processing

Basically these databases for examination score and performance evaluation is two separate databases. So in order to get necessary data for the regression I extracted all the data which includes performance evaluation data and then found the names of people who taken and entrance examination between 2009 and 2012. The basic database for the regression was a table including the data comes from above two different databases. Items included in dataset is : Citizenship registration number, Last name, First name, Gender, Birthdate, Ethnicity, Social class, Position, Position classification, Position ranking, Position date, Position Degree, Performance evaluations scores, Performance evaluation year, years of Education, Entrance examination scores, Date of examination, Organization code, Organization name, Organization size, Status of organization, Self-evaluations for IT, Language and other types of skill, and awards received.

4.3.4. Considerations about the Validity of Data

Correctness of the data inputted to HRMS has been paid special attention before running the regression because author has experienced some incorrect inputs in the database during his work experience in CSC. When collecting the data to HRMS from the government organizations as a group or individual organization, CSC gave some priorities in the data to be checked about the report since the data collected is so vast even for individual worker and covering all the workers in the organization. During the period of 2009-2012,the performance evaluation data was not included in that priority to

input all the data and checked when receiving reports. As a result it has been observed some missing data or incorrect data for the performance evaluation data. Since the HRMS data report is inputted in most of organization by the human resource management officer of the organization different workloads of data input is expected for them depending on the size of the organization. Possibly for this reason there have been detected some inaccuracy in the included in reports and corrected during the period 2009 to 2012.In data preparation process similar possible inaccuracy have been checked and data have removed from base database whenever there is a suspicion. For performance evaluation data the sum of 3 components of evaluation did not meet the letter that should assigned to them as explained in "Performance evaluation in government organization" section above 0-59 should market as "E", 60-69 should marked as "D", 70-79 should marked as "C", 80-89 should marked as "B", and 90-100 should marked as "A".For example, when the sum of 3 components evaluation is 82 and it should be assigned as "B" by letter there were some cases that the letter inputted by the HRM officials as "C" or "A". The datasets which has more than 2 level difference in inputted data and letter should be assigned according to regulation has removed. Datasets with +,-1 difference is left since it might be result of miss click because the evaluation of 3 components as number is actually has to be inputted from keyboard while letter that must be assigned to them is chosen from the choices from putdown arrow. Among final data N=3697, 394cases were -, + 1 level of difference and the rest 3303 cases had no difference. In total from the base database of 6311 datasets 4721 datasets left after checking the

validity of data. Among this 4721 datasets there were some cases that performance evaluation is conducted precedent to civil servants entrance examination. This kind of case happens in practice in Mongolia because some people enter the government without taking entrance examination. There are some ways that people enter government without taking examination. For example, some people have appointed as a temporary basis but worked as temporary until they pass the entrance examination. However, in this thesis since main dependent variable is evaluation score and main independent variable is examination scores here focused on cases in that examination is precedent to performance evaluation. 3697 cases met this condition and used for multiple regression analysis.

4.4. Research Model

Various studies used many different models and third variables when they evaluate validity of the predictors for the job performance. Base model of this study is:

$Perf = \alpha + \beta_1 * Exam + \beta_2 * Age + \beta_3 * GenderD + \beta_4 * Education + \beta_5 * PosRank + \beta_6 * OrgSize + \beta_7 * YDiff + \beta_8 * OrgTypeD1 + \beta_9 * OrgTypeD2 + \beta_{10} * OrgTypeD3 + \beta_{11} * OrgTypeD4 + \beta_{12} * PosDeg + \beta_{13} * IT + \beta_{14} * Lang + \beta_{15} * AwardD + \beta_{16} * ExpM + \varepsilon$

Where:
- Perf = Job performance evaluation scores by supervisors. Range is 0 to 100,

- Exam = Sum of civil service entrance examination score. Range is 55 to 100,

- Age = Age of employee as of examination date,

- GenderD = Dummy variable for gender. 0 – for female, 1 – for male,

- Education = Education calculated by year,

- PosRank = Rank of the Position of employee. Takes a value 1 to 4,

- OrgSize = Number of employees of the organization for the performance evaluation is conducted,

- YDiff = Year Difference between evaluation year and examination year + 1,

- OrgTypeD1 to D4 = Organization Type Dummy variable1 to 4. Baseline for this dummy variable is public organizations which located in primary level of administrative unit, which is called "Sum" and "Duureg". D1 is for organizations located in secondary level of administrative unit, which is called "Aimag" and "Niislel". D2 is for "Agency" which is part of central government located under the ministries in government structure. D3 is for Ministries and Organization under Parliament of Mongolia. D4 is for the highest level of organizations such as office of cabinet and parliament etc.

- PosDeg = Degree of the position. (explained below),

- IT = Sum of answers to self-evaluation questions related to Information Technology knowledge in HRMS,

- Lang = Sum of answers to self-evaluation questions related to Language literacy in HRMS. Questions asked to evaluate language skill in 4 categories such as speaking, listening, reading, and writing. Speaking and writing evaluations multiplied by 2 while listening and reading evaluations multiplied by 0.5 to increase the weight of relatively valued language skill in practice,

- AwardD = Dummy variable for awards earned, 1 - If the employee earned some kind of public award, 0 – If none,

- ExpM = Employee's experience calculated by month as of the date of performance evaluation,

- ε = Error term

In order to see which component is most relevant to job performance another model where components of entrance examination is taken as an individual variable instead of taken as a sum is checked.

4.4.1. Notes about Third variables

Age: Since there is a difference between year of performance evaluation and year of examination age can be calculated differently. Age is taken as of date of examination date. Since this study focuses on the entrance examination, data for the people whose age is over 40 is removed (around 1460 datasets) in very early stages of data processing. In Mongolian public sector human resource management system, administrative and special civil service positions are required to take civil service examination. There are some cases which occur due to the change of the position type, and some public servants

are required to take an examination while they are working in their organization not as an entrance examination. Those people are expected to be the people who took an examination at age over 40.

YDiff: Schmidt and Hunter (1998) showed there is strong learning effect that contributes to the difference in performance between high and low ability people. But it is expected that learning takes some time. New employee might need to adapt to the new environment, to learn most basic skill that all can learn easily or just to get to know partners to work efficiently later. To test this effect YDiff variable is introduced that measures number of years between taking entrance examination and performance evaluation.

Position Ranking: Under the circumstances where performance evaluation is not directly related to salary increment or promotion and performance evaluation is generally viewed as a somewhat symbolic by the public, people in a high ranking position might get higher evaluation than lower ranking officials. To control this effect PosRank variable is included.

Organization Size: Organization size also might affect performance evaluation. Since there is no quota for the rating scales such as "A" evaluation must be under 10% as an example, in the smaller organization where there competition between workers is low the supervisor might not want to demotivate people by giving them symbolic "bad" evaluations. That is why bigger organizations are expected to have relatively lower performance evaluation in average than smaller organizations. For this reason OrgSize variable is included.

Organization Type(Status): It is expected that employees in higher status of government organization have passed higher requirements so they would have higher ability and generally would have higher performance if the performance evaluation truly reflects their ability. To test whether this expectation is true OrdType Dummy variables are introduced. Assuming that higher the organization status the better the performance evaluation.

Position Degree: Parliament Resolution 73 of 1995 regulates position degree and its salary increment of public officials. Each of 5 ranking has 4 degrees, in total there are 20 possible combinations. 10 to 35 percent increase in monthly salary is granted depending on the ranking and degree given. Conditions to award degree to an employee are based on the years of experience as public official, years of experience on relevant position, and performance evaluation. If the performance evaluation process is not valid and more symbolic the real performance that is not measured by performance evaluation should be appear whether employee is given degree or not. However there needs some assumptions such as checking procedure to award degree is more valid than performance evaluation process and supervisors are not reflecting true performance that they observe in the performance evaluation. Under these assumptions people with degree should perform better than those who did not given degree. Or just to test that the position degree which comes with monthly salary increase is truly being given to the right people, PosDeg variable is included.

IT literacy: IT literacy of the employee might affect performance evaluation because in daily work they can help other employees and supervisors when they face IT issues, or they might specifically assigned to tasks that require IT skills and that behavior might affect performance evaluation. However, in the performance evaluation process these kinds of behavior should not be included because those people are not IT specialist [4] they have other tasks that is not IT work according to their responsibilities set by job description. To control this effect variable IT is included in regression model.

Language literacy: Language literacy has also similar logic with IT knowledge. People with more language literacy might have higher performance evaluation. To control this effect variable Lang is included.

Awards: People earned a public award [5] might have higher performance than others. This variable is included as it might reflect the persons' general ability that should affect both examination score and performance evaluation. If the other variables that should be valid for job performance found insignificant and award variable is found significant we can conclude that performance evaluation process is not reflecting job performance but just being done symbolically.

---

[4] IT specialists are not included in this study because they are not required to take an examination.
[5] Usually awards that related to their job performance is included.

Chapter V. ANALYSIS

5.1. Results and Analysis

5.1.1. Descriptive statistical analysis

Descriptive statistics of the data used in this study is shown below table. 3697 datasets have used in a model. Forty percent of the subjects were male. The average ages of subjects were 28.07.Subjects had 16 years of experience in average. Subjects had22.5 months of job experience on average. In average subjects had near 2 years of experience.[6]However there were some extreme cases according to the length of experience. Average score of performance evaluation of subjects were 85.68, and average score of examination score were 67.60 out of 100. Average size of the relevant organizations was93 workers. However, there were some extreme organizations which is very small or very big.

Table 3. Simple Statistics

Variable	N	Mean	Std Dev	Min	Max
Perf	3697	85.68	7.62	35	100
Exam	3697	67.60	7.07	55	87.49
Age	3697	28.07	5.03	20	40
GenderDD	3697	0.40	0.49	0	1

---

[6]123 subjects had more than 5 years of experience the rest was under 5 years and 23 subjects had an experience less than month. 49.8 % of the subjects located inside IQR.

Education	3697	16.01	0.94	12	21
PosRank	3697	1.75	0.56	1	4
Log_OrgSize	3697	3.95	1.03	0	6.36
OrgSize	3697	92.74	129.52	2	579
YDiff	3697	2.09	0.95	1	4
OrgTypeD1	3697	0.39	0.49	0	1
OrgTypeD2	3697	0.09	0.29	0	1
OrgTypeD3	3697	0.06	0.24	0	1
OrgTypeD4	3697	0.03	0.16	0	1
PosDeg	3697	1.45	1.89	1	15
IT	3697	13.83	5.88	2	36
Log_Lang	3697	2.58	0.51	1.61	3.92
AwardD	3697	0.05	0.21	0	1
Log_ExpM	3697	2.79	0.89	-1.32	4.97
ExpM	3697	22.49	17.51	0.27	143.9

5.1.2. Correlation Analysis

From the correlation matrix below we can see that all the independent variables are significantly correlated to dependent variable except dummy variables. Only PosDeg and Log_OrgSize variables are positively correlated with performance evaluation at .05 significant level. All other variables are correlation with dependent variable at .01 significant level. Furthermore, no strong correlation between

independent variables is detected. Highest correlation calculated is between YDiff and ExpM variables. This is natural because these two variables are both related to experience. However, for explaining job performance and from the model it is considered bearable to keep these two variables.

Table 4. Correlation Table

| | Pearson Correlation Coefficients, N = 3697 | | | | | | | | | | | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1)Perf																
(2)Exam	.07***															
(3)Age	.05***	.03*														
(4)GenderDD	.00	-.01														
(5)Education	.06***	.06***	.11***													
(6)PosRank	.18***	.05***	.18***	.08***												
(7)Log_OrgSize	.04**	-.01	.15***	-.03*	.08***											
(8)YDiff	.16***	-.02	.03**	.00	.06***	.07***										
(9)OrgTypeD1	.07***	.00	.04**	.02	.07***	.2***	.00									
(10)OrgTypeD2	.00	-0.02	.07***	.08***	.11***	.17***	.14***	.00								
(11)OrgTypeD3	.02	.07***	.04**	.06***	.14***	.26***	.09***	.03**	-.21***							
(12)OrgTypeD4	.14***	.03**	.03*	-.01	.06***	.31***	.26***	.03*	-.13***	-.05***						
(13)PosDeg	.03**	.04**	.07***	-.03**	-.02	.05***	-.1***	.02	-.02	-.02	.1***					
(14)IT	.11***	.07***	-.06***	.15***	.21***	.28***	.1***	.03*	.02	.11***	.2***	.05***				
(15)Log_Lang	.1***	.03*	.15***	.03	.15***	.19***	.18***	.05***	-.02	.08***	.18***	.09***	.05***			
(16)AwardD	.07***	.04**	.08***	.03**	.07***	.12***	-.03	.04***	.06***	.07***	-.02	-.01	.02	.11***		
(17)Log_ExpM	.19***	.01	.12***	-.01	-.01	.12***	.04**	.48***	.07***	.00	-.03	.08***	.02	.01	.06***	.09***

*** $p < 0.01$;
** $p < 0.05$;
* $p < 0.1$

Normality assumption of OLS regression method has been checked. Variables OrgSize, Lang, and ExpM have used after taking log to improve normality assumption. Some concerns remain about the normality of residuals. However, the issue might be caused from characteristics of some variables such as Perf, Total, and Age. Perf and Total have a upper limit while Age had practically both lower and upper limit. Truncated regression has also been run and results compared to OLS method. No big difference observed. Based on these considerations normality issue is considered negligible and OLS method is used.

Multicollinearity issue has also checked. VIFs for variables ranged from 1.01 to 1.71. For all variables VIF value were smaller than 10. However, considering that relatively small R-Square value collinearity diagnostics with Eigenvalue and Condition index has also been done. There were no Eigenvalue that is close to 0 detected. For Condition Index no value was greater than 30. Multicollinearity has not detected from these values too.

5.1.3. Regression Analysis

Regression results are as follows:

Table 5. Regression Results

	Base Model				Component Model		
Variable	**Para Est**	**S. E**	**Pr > |t|**	**Variable**	**Para Est**	**S. E**	**Pr > |t|**
Intercept	71.943	2.472	<.0001	**Intercept**	73.216	2.59	<.0001
Exam	0.059	0.017	0.0006	**Multiple**	0.031	0.02	0.1196

				Doc	0.284	0.119	0.0171
				Interv	0.036	0.046	0.4243
				Competence	0.18	0.047	0.0001
Age	0.01	0.025	0.6879	Age	0	0.026	0.987
GenderDD	-0.195	0.25	0.435	EducationYear	-0.067	0.142	0.6363
Education	0.008	0.135	0.9537	GenderDD	-0.174	0.251	0.4889
PosRank	1.478	0.281	<.0001	PosRank	1.475	0.281	<.0001
Log_OrgSize	-0.21	0.128	0.1	Log_OrgSize	-0.206	0.128	0.1074
YDiff	0.658	0.146	<.0001	YDiff	0.633	0.146	<.0001
OrgTypeD1	0.537	0.3	0.0738	OrgTypeD1	0.538	0.3	0.0732
OrgTypeD2	-0.443	0.489	0.3646	OrgTypeD2	-0.44	0.491	0.3709
OrgTypeD3	-0.791	0.597	0.1853	OrgTypeD3	-0.956	0.606	0.1146
OrgTypeD4	4.308	0.885	<.0001	OrgTypeD4	4.258	0.885	<.0001
PosDeg	0.068	0.065	0.2939	PosDeg	0.071	0.065	0.2722
IT	0.082	0.023	0.0004	IT	0.08	0.023	0.0005
Log_Lang	0.703	0.259	0.0067	Log_Lang	0.66	0.26	0.0113
AwardD	0.955	0.581	0.1006	AwardD	0.959	0.582	0.0997
Log_ExpM	1.058	0.157	<.0001	Log_ExpM	1.06	0.157	<.0001
F value	22.25***			F value	19.41***		
R-Square	0.0882			R-Square	0.0912		

5.2. Interpretation of Results

Table 5 reports the results of model explained in section 4.4. The second column reports the results of the base model in which the independent variable is sum of examination scores. In component model, components of entrance examination are taken as a separate variable.

51

The results of base model shows that sum of examination scores (β=0.059, p=0.0006) was significantly related to performance evaluation score. People showed higher performance on entrance examination performed better after they are employed to the public organizations. The main hypothesis of this study has accepted. There is a positive relationship between interested variables however absolute value of the parameter is not so high. Sum of examination scores along with the control variables accounted for 8.8% of the total variance in the leniency of ratings (R^2= 0.0882, p<0.001).After removing outliers with dffit criteria that is less than 2*sqrt(k/n) where k is number of predictors and n is number of observations, R-square of model increased to 0.13 and βincreased to 0.065. After removing outliers component model results also increased. Multiple choice test examination became significantly related to job performance (β=0.043, p=0.0187) .

Age can affect both the performance evaluation and examination score. As age increases candidates may have more experience of taking various types of examination and it may increase their score on examination. On the other hand as age increases employees attain more experience of job or job evaluation itself or other factors that may affect such as how to deal with communication issues with supervisor or daily hardship. Age and performance evaluation score would have a positive relationship. However no significant relationship between age and performance evaluation even though the sign was positive (β=0.010). This result is consistent with the results of Schmidt and Hunter (1998) and previous studies. There were positive

correlation between age and examination score at 0.1 significant level (r=0.03*).

Education also could have similar relationship with examination score and performance evaluation. Education is expected to have positive relationship with both these variables. The results show that the education measured as years of education was not significantly related to performance evaluation (β=0.008, p=0.9537). There were slight correlation between education and examination score (r=0.06***).However, this might be resulted from validity of data concerning education.

The results reported that higher ranking officials are getting higher performance evaluation score(r=1.48, p <0.001).Although this could be resulted from the difference in the ability it is also consistent with the expectation that higher ranking officials might get higher evaluation score under performance evaluation is somewhat symbolic. Further investigation is required to clear the cause.

Expectations about the organization size effect to the performance evaluation were confirmed from the results. In smaller organizations where competition among employees is low, the supervisor might not want to demotivate people by giving them symbolic "bad" evaluations. Though significance level is 0.1 there were negative relationship between organization size and performance evaluation. (Parameter was estimated as: β = -0.00239 were, p =0.025 in a model without taking log from orgsize).

Schmidt and Hunter (1998) showed there is strong learning effect that contributes to the difference in performance between high and low ability people. But it is expected that learning takes some time. To test this effect average performance evaluation scores of all employees were calculated.

Table 6.

Year	0	1	2	3
Avg Perf	84.43	85.56	86.45	88.86

As year count after the employee passed entrance examination increases average performance evaluation score increased. This is consistent with the expectation. The results reports also consistency with the expectation (β =0.658, p <0.001).

It is expected that employees in higher status of government organization have passed higher requirements so would have higher ability and generally would have higher performance if the performance evaluation truly reflects their ability. Results reported that compared to lowest level of organizations regional organizations has 54% higher performance in general (β =0.537, p =0.0738) and highest level of organizations has 4.31 times higher performance (β =4.308, p <0.001). However, agency and ministry level organizations had estimated parameters of -0.443 and -0.791 with p0.365 and 0.185 respectively. These results were opposite of the expectation, however this might be the caused with another variable, organization size because agencies and ministries are relatively large in organization size. Further investigation is required to reveal the cause.

IT and language literacy both seems to influence performance of the employee. Parameters estimated were 0.082 (p =0.0004) and 0.703 (p =0.0067) both statistically significant. It seems that aside from skill and knowledge to perform main responsibilities for job, IT and language skills are required for young generation in current Mongolian government organizations.

The results shows that position degree, supposedly should be based on the outstanding performance of the employee, has no statistically significant relationship with performance evaluation (β =0.065, p = 0.2939). Sign of the relationship is positive though. Further research is required to test this result.

Experience has been found a valid predictor for performance with r=0.18 according to Schmidt and Hunter (1998). However, Schmidt, Hunter, and Outerbridge (1986) found that when experience on the job does not exceed 5 years validity for job experience was .33. Job experience exceeding 5 years did not result increase in job performance. Since most of the subject had experience less than 5 years it was impossible to test this results. Results report that log_expM related to performance significantly (β =0.157, p = 0.0001) where parameter for expM is estimated 0.0549 with 0.0001 probability for subjects under 5 years of experience.

Overall, 2 models show similar results. Among the components of entrance examination, competence and document examination score showed strongest predictor β =0.284, β =0.18 respectively for job performance.

Individual interview and multiple choice examination score did not show significant relationship with job performance. Hypothesis 2 assuming that the multiple choice test examination will be the strongest predictor for job performance is not accepted.

Moreover, since that the expectations of the variables and hypothesis were found consistent with the results in above table, it is concluded that performance evaluation is not just symbolic as believed by public.

5.3. Qualitative background

Small interview has been conducted to focused group, who has wide experience and knowledge about human resource management, in order to check the regression results on practice. The human resource management officials of government organizations who kept their position after 2012 election and considered to possess sufficient knowledge about their organizations' human resource management were included in the interview. Questions to compare job performance of people with high examination score and low examination score, to evaluate employees' competency related to components of examination were asked.

In average interviewees were working in their organization 9.5 years. 50.9 people have appointed to relevant organizations during 2009-2012, among them 28.7 people were appointed to the position by civil service entrance examination. 7 interviewees answered that the examination score is related to job performance in general or to some of the employees, among

them 1 interviewee answered the relationship is true in general. For the question to evaluate this relationship from 1 to 3 scale, average response was 2.0. To the question to evaluate average satisfaction level of new employees to organizations' requirement, average response was 3.1 out of 5 point scale. For the questions to evaluate these requirements in detail education and computer skills are evaluated 4.5 and 4.0 out of 5 point scale, while attitude and behavior to work and competence of official documentation standard is evaluated least as 2.9 and 2.5 out of 5 point scale. Knowledge of related laws and regulation is evaluated 3.3 on average.

Chapter VI. CONCLUSION

6.1. Summary of Research

The validity of Mongolian civil service entrance examination for job performance has been tested using data from Civil Service Council of Mongolia. Data sources are entrance examination database and Human Resource Management System (HRMS) which includes general human resource data of all Mongolian government officials. Multiple linear regression has been chosen to estimate parameters in regression model. Assumptions of multiple linear regression have been checked and it is considered Ordinary Least Squares (OLS) method is acceptable for this research. Main dependent variable was a performance evaluation score of the public servants and main independent variable was their score that they get on civil service entrance examination. Age, education, gender, experience, position ranking and degree of official, type and size of the organization that they are working for, IT and language literacy of official, and awards they received is included in the regression model as a control variable. Two models have tested. Base model took sum of examination score as an independent variable. Alternatively "component model" has checked where each components of the entrance examination is taken as an independent variable.

Two hypotheses were formulated and tested.

Hypothesis 1: Performance evaluation score and entrance examination score have positive relationship. The higher the examination score the higher will be performance evaluation score.

Hypothesis 2: Since the multiple choice examination is taken most seriously by both officials and candidates and its weight in the examination score is highest at 60 out of 100, the score for multiple choice examination would be most strongly related to job performance compared to other components of entrance examination.

Hypothesis 1 is accepted since parameter for the relationship between entrance examination and job performance was found positive and significant at 0.01 level (β =0.059, p = 0.0006). R-square of the model was 0.0882.After removing outliers R-square of model increased to 0.13 and β increased to 0.065. After removing outliers multiple choice test examination became significantly related to job performance (β=0.043, p=0.0187) in component model.

Hypothesis 2 is not accepted because multiple choice test examination score did not show significant relationship with job performance level (β =0.031, p = 0.1196). Among the components of entrance examination, document examination score showed strongest predictor (β =0.284, p = 0.0001) and competence examination was estimated second strongest predictor (β =0.18, p = 0.0171) for job performance. Individual interview also did not show significant relationship with job performance (β =0.036, p = 0.4243).

No significant relationship between age and performance evaluation even though the sign was positive (β=0.010, p =0.6879). There were positive correlation between age and examination score at 0.1 significant level (r=0.03*).

Education measured as years of education was not significantly related to performance evaluation (β=0.008, p=0.9537). There were slight correlation between education and examination score (r=0.06***).

The results reported that higher ranking officials are getting higher performance evaluation score (r=1.48, p <0.001).

Though significance level is 0.1 there were negative relationship between organization size and performance evaluation (β=-0.21, p=0.1). (Parameter was estimated as: β = -0.00239 were, p =0.025 in a model without taking log from orgsize).

As year count after the employee passed entrance examination increases average performance evaluation score increased. This is consistent with the expectation. The results reports also consistency with the expectation (β =0.658, p <0.001).

Compared to lowest level of organizations regional organizations has 54% higher performance in general (β =0.537, p =0.0738) and highest level of organizations has 4.31 times higher performance (β =4.308, p <0.001).

However, agency and ministry level organizations had estimated parameters of -0.443 and -0.791 with p0.365 and 0.185 respectively.

IT and language literacy both seems to influence performance of the employee. Parameters estimated were 0.082 (p =0.0004) and 0.703 (p =0.0067) both statistically significant.

Position degree, supposedly should be based on the outstanding performance of the employee, has no statistically significant relationship with performance evaluation (β =0.065, p = 0.2939).

Log_expM was found significantly related to job performance (β =0.157, p = 0.0001) where parameter for expM is estimated 0.0549 with 0.0001 probability for subjects under 5 years of experience.

Small interview has been conducted to focused group, who has wide experience and knowledge about human resource management, in order to check the regression results on practice. Seven out of eight interviewees replied that there is a relationship between examination score and job performance according to their observation on the job.

Moreover, since that the expectations of the variables and hypothesis were found consistent with the results in above table, it is concluded that performance evaluation is not just symbolic as believed by public.

Results of this study were consistent with previous study and findings about the control variables were consistent with the expectations for most variables.

6.2. Discussion and Implication

The results of this study were consistent to previous study (Schmidt and Hunter, 1998). However, compared to values of parameters in meta-analyses, parameter for entrance examination as a predictor for job performance was relatively low (β =0.059, p = 0.0006).While competence examination score was estimated strongest predictor for job performance amongst components of entrance examination, respondents to qualitative survey evaluated competence of new employees for official documentation which is one component of competence examination most unsatisfied. These findings might be result few different conditions in reality. This could be result of general tendency of low competency for official documentation standard which leads to dissatisfaction of competency for official documentation standard, and people with high ability tend to possess that competency. That is why competence examination is estimated strongest predictor for job performance. Or this could reflect that another component of the competence examination, knowledge for computer, is the strongest predictor. However, that seems not fully true because the results of main model shows parameter estimate for IT skill self-evaluation is 0.08**. Further research is required to decide which reality that these results reflect.

Although there needs more consideration about the nature of relevant position and organizational goal etc when deciding what competencies does civil service entrance examination should measure for candidates, it seems more aspects of general cognitive ability should be introduced to the entrance examination to improve its power to predict future performance. Moreover, it became clear that current individual interview examination needs to be revised and renewed in order to increase its validity for job performance. Some considerations are needed to improve weight of stronger predictors and decreasing weight of weak predictors.

6.2.1. Limitations of Study

In this study variables such as sex of rater, rater's opportunity or length of time to observe rate's job performance, rater's supervisory experience, whether in the organization rater and the ratee are meeting and discussing about the performance evaluation or not, and effect of other employees to the subject performance evaluation as an effect of workplace deviant behavior or organizational citizenship did not controlled for the regression model. Due to the frequent change in the workforce in government as a result of still persisting strong political and personal influence to the civil service, determining the rater of the subject was difficult since there was a no data who was the rater. Further research can be done controlling those effects related to the rater or organizational behavior etc that just mentioned above. From the database that used in this study theoretically it was possible to determine the rater, however frequent change in the personnel and lack of

clear connection and sometimes lack of input in the experience history of the personnel that could help who was working on a same period in a same organization and might were the rater, made it difficult to determine the rater. Moreover, further research can be done using robust regression models which deal with some issues found about the assumptions of the OLS method or research to explain detailed reasons of the findings of this research.

Acknowledgements

This research was possible with the contributions from many people directly or indirectly, for whom I would like to express my sincere gratitude.

I would like to appreciate Korean government, and Korean International Cooperation Agency and Seoul National University for giving me this opportunity to study in Korea as a scholarship student.

I also would like to express my deep gratitude to GMPA office, especially Ms.Lee for her support during whole period of our stay in Korea. I am also indebted to all the professors who thought me general knowledge and statistical skills necessary for this study.

I really appreciate my thesis advisor Professor Eom Seok Jin for his valuable comments and guidance. This thesis would not have been possible without hid guidance.

Furthermore, I would like to thank my colleagues in my office at Civil Service Council Mongolia, Mr. Baatarsaikhan, Mr.Sosorbaram, and Mr. Altanbagana for offering timely support and data used in this study. As well, I thank to all the respondents to the interview who have willingly shared their precious time and offering honest opinions.

직무성과를 위한 몽골 공무원 채용시험의 타당성에 관한 연구

AmartuvshinAmarmend

행정대학원 행정학 전공
서울대학교

본 연구는 몽골의 공무원채용시험이 직무성과에 미치는 영향을 살펴보기 위해서 몽골의 인사위원회 자료를 이용하여 분석해보았다. 분석을 위해서 공무원채용시험 데이터베이스와 모든 몽골 공무원들의 인사자료를 포함하고 있는 인적자원관리시스템(HRMS)을 활용하였다. 다중회귀분석을 실시하였으며 특히 OLS 회귀분석을 통해서 가설을 검증하였다. 주요 종속변수는 공무원들의 업무평가 점수이고, 독립변수는 그들이 공무원채용시험에서 받은 점수였다. 그 밖에 나이, 교육수준, 성별, 경력, 직급, 직위, 그들이 일하는 조직의 유형과 크기, IT 기술, 어학실력 그리고 보상수준 등을 통제변수로 설정하였다.

연구 결과, 몽골의 공무원채용시험 성적과 직무성과는 통계적으로 유의미한 수준에서 관계가 있는 것으로 나타났다(β =0.059, $p<0.01$). 채용시험 중에서도 서류전형 점수가 업무성과와 강한 관계가 있는 것으로 나타났으며(β =0.284,$p<0.01$), 역량평가 점수 역시 두 번째로 강한 영향력이 있는 것으로 나타났다(β =0.18,

p<0.01).다중문항선택 시험의 점수는 통계적으로 유의한 관계가 나타나지 않았다. 개별 면접의 경우에도 업무성과와 관계가 없는 것으로 나타났다.

본 연구의 결과는 포커스그룹 인터뷰(FGI)를 통해 확인한 결과와 일치하였다. 결론에서는 향후 공무원채용시험 발전을 위한 정책적 함의를 제시하였다.

Bibliography

Antonioni, D. and Park, H. (2001). "The relationship between rater affect and three sources of 360-degree feedback ratings." *Journal of Management,*27, 479–495.

Beilock, S. L. and M. S. DeCaro. (2007). "From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure." *Journal of Experimental Psychology-Learning Memory and Cognition,*33(6), 983-998.

Bello-Imam, I. B., Oshionebo, B. O. and Ojeifo, S. A. (2007). *Fundamentals of Human Resource Management in Nigeria.* Ibadan, College Press and Publishers Ltd.

Bommer, W.H., Johnson, J., Rich, G.A., Podsakoff, P.M., et al. (1995). "On the interchangeability of objective and subjective measures of employee performance: A meta-analysis." *Personnel Psychology*, 48, 587-605.

Borman, W. C. and Motowidlo, S. J. (1993). "Expanding the criterion domain to include elements of contextual performance." In N. Schmitt & W.C. Borman and Associates (Eds.), *Personnel selection in organizations,* 71-98. San Francisco, CA: Jossey-Bass Publishers.

Borman, W. C., White, L. A., Pulakos, E. D. and Oppler, S. H. (1991). "Models evaluating the effects of ratee ability, knowledge, proficiency, temperament, awards, and problem behavior on supervisory ratings." *Journal of Applied Psychology, 76,* 863-872.

Campbell, J. P. (1990). "Modeling the performance prediction problem in industrial and organizational psychology". In M. D. Dunnette and L. M. Hough (Eds.), *Handbook of industrial and organizational psychology*(1),687-732. Palo Alto, CA: Consulting Psychologists Press.

Crino, M. D. (1994). "Employee sabotage: A random or preventable phenomenon." *Journal of Managerial Issues,* 6(3)*,* 311-330.

Cronbach, C. J. (1970). *Essentials of Psychological Testing.* New York: McGraw – Hill

DeCaro, M. S., et al. (2008). "Individual differences in category learning: Sometimes less working memory capacity is better than more." *Cognition,*107(1), 284-294.

Donald E.Klinger, John Nalbandian, and Jare Llorens.(2010). *Public Personnel Management: Contexts and Strategies.* 6th edition,Longman.

Dunnette, M. D. (1972). *Validity study results for jobs relevant to the petroleum refining industry.* Washington, DC: American Petroleum Institute.

Dunlop, P.D., and Lee, K. (2004). "Workplace deviance, organizational citizenship behavior, and business unit performance: the bad apples do spoil the whole barrel." *Journal of Organizational Behavior* 25, 67–80.

Farh, J.-L., Werbel, J.D., and Bedeian, A.G. (1988). "An empirical investigation of self-appraisal-based performance evaluation." *Personnel Psychology*, 41, 141-156.

Frank L. Schmidt and John E. Hunter.(1981). *American Psychologist*, 36(10), 1128-1137.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26,* 461-477.

Ghiselli, E. E. Campbell, J. P. and Zedeck, S. (1981). *Measurement theory from the Behavioral Sciences.* San Francisco. Freeman.

Giacalone, R. A., and Greenberg, J. (1997). *Antisocial Behavior in Organizations.* Thousand Oaks, CA: Sage.

Gimmig, D., et al. (2006). "Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence." *Psychonomic Bulletin and Review,*13(6), 1005-1010.

Glass, G. V. (1976). "Primary, secondary, and meta-analysis of research". *Educational Researcher, 5,* 3-8.

Glen Grimsley and Hilton F. Jarrett.(1973). "The relation of past managerial achievement to test measures obtained in the employment situation: Methodology and results". *Personnel Psychology,*26, 31-48.

Glen Grimsley and Hilton F. Jarrett (1973). "The relation of past managerial achievement to test measures obtained in the employment situation: Methodology and results - II". *Personnel Psychology,*28, 215-231.

Hunt, S. T. (1996). "Generic work behavior: An investigation into the dimensions of entry-level, hourly job performance." *Personnel Psychology, 49*, 51-83.

Hunter JE, and Hirsh HR. (1987). "Applications of meta-analysis." *International review of industrial and organizationalpsychology,* 321-357. Chichester: Wiley.

Hunter JE, and Hunter RF. (1984). "The validity and utility of alternative predictors of job performance." *Psychological Bulletin*, 96, 72-98.

Hunter, *S.* E., Schmidt, F. L., and Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Kahya, Emin.(2007). "The effects of job characteristics and working conditions on job performance", *International Journal of Industrial Economics*, 37(6), June 2007, Pages 515–523.

Katz, D., and Kahn, R. L. (1978). *The social psychology of organizations.* New York: Wiley.

Lang, J. W. B. and P. D. Bliese. (2009). "General Mental Ability and Two Types of Adaptation to Unforeseen Change: Applying Discontinuous Growth Models to the Task-Change Paradigm." *Journal of Applied Psychology,94*(2). 411-428.

Lautenschlager, G. J., and Shaffer, G. S. (1987). "Reexamining the component stability of Owen's biographical questionnaire". *Journal of Applied Psychology*, 72(1), 149-152.

Lee, D.M., and Alvares, K.M. (1977). "Effects of sex on descriptions and evaluations of supervisory behavior in a simulated industrial setting." *Journal of Applied Psychology,*62, 405–410.

McEnery, J.M., and Blanchard, P.N. (1999). "Validity of multiple ratings of business student performance in a management situation." *Human Resource Development Quarterly,* 10, 155-172.

Michael Armstrong.(2001), *A Handbook of Human Resource Management Practice,* 8th edition*,* Kogan Page.

Moser, K., Schuler, H., and Funke, U. (1999). "The moderating effect of raters' opportunities to observe ratees' job performance on the validity of an assessment center." *International Journal of Selection and Assessment,*7 (3), 133–141.

Murphy, K. R. (1989). "Dimensions of job performance." In Dillon R, Pellingrino J (Eds.), *Testing: Applied and theoretical perspectives,* 218-247. New York: Praeger.

Norma M.Riccucci, Katherine C.Naff with assistance of Jay M.Shafritz, David H.Rosenbloom. (2008). *Personnel Management in Government Politics and Process.* CRC Press.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome.* Lexington, MA: Lexington Books.

Reilly, R. R., and Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology, 35,* 1-62.

Rotundo, M. (2000). "The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *"Dissertation Abstracts International, 61*(5-A), 2058.

Schmidt, F. L. (1992). "What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology." *American Psychologist, 47,* 1173-1181.

Schmidt, F. L., and Hunter, J. E. (1977). "Development of a general solution to the problem of validity generalization." *Journal of Applied Psychology, 62,* 529-540.

Schmidt, F. L., and Hunter, J. E. (1998)."The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *"Psychological Bulletin*, 124, 262-274.

Schmidt, F. L., Hunter, J. E., and Outerbridge, A. N. (1986). "The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance." *Journal of Applied Psychology, 71,* 432-439.

Schmidt, F. L., Hunter, J. E., Pearlman, K., and Shane, G. S. (1979). "Further tests of the Schmidt-Hunter Bayesian Validity Generalization Model." *Personnel Psychology, 32,* 257-281.

Schmidt F.L, Ones D.O, and Hunter J.E. (1992). "Personnel selection." *Annual Review of Psychology, 43,* 627-670.

Vineberg, R., and Joyner, J. N. (1982). *Prediction of job performance: Review of military studies.* Alexandria, VA: Human Resources Research Organization.

Wallace B. Hall and Harrison G. Gough.(1977). "Selecting statistical clerks with the Minnesota Clerical test", *The Journal of Psychology*, 1977, 96, 297-301.

Wayne F.Cascio, and Herman Aguinis, (2005), *Applied Psychology in Human Resource Management,* 6th edition, Pearson Prentice Hall.

Witt, L.A., Kacmar, K.M., Carlson, D.S., and Zivnuska, S. (2002). "Interactive effects of personality and organizational politics on contextual performance.*" Journal of Organizational Behavior, 23, 911–926.*

Yun, G.J., Donahue, L.M., Dudley, N.M., and McFarland, L.A. (2005). "Rater personality, rating format, and social context: implications for performance appraisal ratings." *International Journal of Selection and Assessment*, 13 (2), 97–107.