Quality of Supervision of Ph.D. Program among Public Universities in Malaysia: A Rasch Model Analysis

Mikail Ibrahim[1], Siti Aishah Hassan[2]

[1] Faculty of Major Language Studies, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800, Nilai, Negeri Sembilan, Malaysia.
[2] Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
siti_aishahh@putra.upm.edu.my

Abstract: This study examines Ph.D. students' satisfaction with the supervision process at four selected universities in Malaysia. In addition, the study also investigated the psychometric properties of Quality Supervision Scale (QSS); specifically the scale dimensionality, construct validity, endorsibility, and estimation of item and person score reliability of the scales. The participants were 153 Ph.D. students of these universities. The QSS includes many qualities of effective supervision such as supervisor academic competency, research methods competency, attitude towards supervisee, faculty academic and moral supports and supervisees' personal traits was distributed to the respondents. The Rasch model analysis was employed to analyze the data for reliability, fit to the model, estimation of satisfaction levels and possibility of scale to function differentially across gender. Results suggested that generally students were satisfied with the supervision processes at these universities. In addition to that, the scale satisfied psychometrics properties by maintaining unidimensionality, reliability, and internal consistency. Furthermore, Rasch analysis revealed that, for gender, differences in overall satisfaction levels between males and females were marginal. The differential item functioning showed that only 6 of 49 calibrated items function differently. This suggested that students' levels of satisfaction were constant across gender. However, the study recommended that future studies should examine the satisfaction level across different disciplines since previous studies suggested that satisfaction differs across different domains.
[Mikail Ibrahim; Siti Aishah Hassan. Faculty of Major Language Studies, Universiti Sains Islam Malaysia; Faculty of Educational Studies, Universiti Putra Malaysia. Journal of American Science 2011; 7(2):562-575]. (ISSN: 1545-1003). http://www.americanscience.org.

Keywords: supervision, post-graduate studies, Ph.D. students, Rasch model

1. Introduction

In the past few decades, postgraduate studies in general and Ph.D. supervision in particular had received massive attention from scholars, practitioners, authorities and stakeholders alike (Chiang, 2003; Brew & Peseta, 2004). The main reason for this enormous attention is to refine, renovate and reenergize the postgraduate studies since the Ph.D. programme faces many challenges. The social science research on doctorial candidates and supervision has never included prolonged and systematic observation of supervision as actually happening in natural sciences, where there are more general laboratory courses, the supervision of postgraduate and post doctorate work has been observed as de facto (Hockey, 1991; 1995).

Researchers (Hockey, 1995, Burgess, 1994; Wright & Cochrane, 2000; Chiang, 2003; Marsh, Rowe, Martin, 2002; Brew & Peseta, 2004) referred the problem of the PhD programme to different opinions and views on what constitute of a Ph.D. programme. There are two major different views of what constitutes of Ph.D. The first view perceives Ph.D. process as a traditional knowledge-based in which candidates would be given a huge quantity of knowledge in various aspects of disciplines within the scope of the specialization to assist their work and to carry out independent research in the future. Conversely, there are some scholars who view Ph.D. process as a large of formal research training in which students would be trained on research without necessarily focusing on specific body of knowledge per se (Hockey, 1991). This standpoint believes that Ph.D. candidates have acquired necessary and sufficient body of knowledge in their previous studies that can enable them to carry out their project successfully. However, they need urgent resejjarch training and wide range of research methods on the way to effectively utilize their knowledge, in order to prepare them for the future research endeavours.

In 1988 in the United Kingdom, the committee of vice-chancellor and principal (CVCP) combined both views and emphasized that Ph.D. process should comprise these two understanding because it would be hard if not impossible for the candidates to successfully complete the programme without having enough knowledge in both theories and methods. Thus, the committee highlighted that the objectives of a Ph.D. programme are as follows:

- The first is to enable young people of high intellectual ability to develop and bring to fruition as far as possible the quality of originality, to contribute new and significant ideas, to make a positive contribution to knowledge and creativity in their respective disciplines.

- The second is to provide a training in research methods which makes them capable subsequently of assuming the role independent scholars and research writers at the highest levels capable of planning and carrying to completion a well-perceived plan of research directed toward a given objective without necessity in supervision from experienced people

Two and half decades ago and precisely in 1980s, higher institutions particularly in UK started to reevaluate the whole postgraduate programmes especially the Ph.D. after it became realistic that the overwhelming majority of Ph.D. candidates could not successfully complete their study within the time frame (Burgess, 1994; Brew & Peseta, 2004). Similarly in Austaralia, the proportion of the research students who fail to submit a thesis after a period of public or industrial supports has been increase (Buttery, Rithcher & Filho, 2005). In other words, low rate submission and successful completion of Ph.D. theses or dissertations had mainly instigated educationists and practitioners to devote their attention to examine the factors that might affect the problems. It was evidenced that, only 18.2 percent of total enrolled students in social science in 1980 finished their study within the allocated time. Consequently, the Economic and Social Research Council (ESRC) enacted a rule to withdraw grants for two years and sanction any institution where the Ph.D. submission is less than 10% in 1985, 25% in 1986, 35% in 198, 40% in 1988 and this percentage had increased until it reached 50% submission rate by 1993 (Burgess, 1994).

Historically, according to Wright and Cochrane (2000), a Ph.D. programme based on research is very few in Britain due to the fact that postgraduate research-based was established in Germany and spread to the United States since the nineteenth century but later adopted by Britain. Moreover, in 1960s, after the establishment of new higher institutions in UK as well as in other Western countries, and after influential Swinnerton-Dyer report (Wright & Cochrane, 2000) about the situation of postgraduate studies especially PhD, studies started to investigate the quality of research and the process of supervision. Two decades ago, Swinnerton-Dyer report recommended a series of suggestion, including: providing students with a grounding in research techniques, taking account of supervision quality, and completion within the time frame, and evaluating students' performance as well as supervisors'. The report also suggested many disciplinary acts against any institution with poor rates of submission and satisfactory completion of the research within time frame, maximum 4 years.

Harris (1996) asserted the importance of the role of supervisor in successfully completing the PhD programme. He recommended a series of guideline to energize the programme as reported in previous reports. The report emphasized on supervision, infrastructure and environment, monitoring and assessment. These reforms were considered to be an evolution to the postgraduate studies especially the Ph.D. in UK. They also transformed the UK postgraduate programme from been a traditional institutions to modern institutions in which students would acquire the necessary knowledge and sufficient research methodologies, also would be able to finish the study within the allocated time and resources.

Studies suggested that supervision is one of the major determinants of Ph.D. program succession whether in terms of efficient and effective information provides to the students in both concept of knowledge and research methodology, or in term of maintaining the student motivation through collegial stimulus and support throughout students' research training (Brew & Peseta, 2004; Seagram, Gould & Pyke, 1998). Hockey (1991) mentioned two major factors to successfully complete of Ph.D. programme; students' ability and motivation. These two factors could be formed by using two strategies; knowledge based and training in research methods. However, Blume (1995) stated that "further scientific work, leading to the title of doctor was not perceived as a training in research, but a research itself" (p.11).

However, Hockey (1991) suggested that the nature of supervision is a very crucial element in determining the completion of a Ph.D. study. In his review of literature of many studies on supervision, he found that generally there is students' dissatisfaction with supervision processes. According to him, a general trend among Ph.D. students indicated that 25% of social science students were dissatisfied with their supervision process and emphasized that they received too little supervision in the initial stage of their research. The study also revealed how bitterly students were in all stages of their studies. On the other hand, Buttery, Rithcher and Filho (2005) opined that the efficiency of Ph. D supervision process depends on various stages of thesis lifecycle. At the early stage, single supervision is considered the best as the students is in the brainstorming and planning of the study. They emphasized the best way to confused the student is by

giving a conflicting advice which may detract from clear focus and research direction. However, when the research direction and questions have been identified then only it would be beneficial to discuss in group supervisions.

Interestingly, studies also suggested that the dissatisfaction of postgraduate students towards the processes of supervision differed across the disciplines (Young, Fogarty, McRae, 1987). The study indicated that social science students rated generally higher than the natural sciences in the dissatisfaction of the supervision process. The reason for the dissatisfaction of social science students might not unrelated to the nature of research in human sciences and characters of supervisors. Unlike natural sciences, social sciences' students expected to establish strong relationships with supervisors to facilitate and solve their academic problems. It was suggested that more attention should be paid to the entire Ph.D. supervision process especially in human science where completion rate and submission was poor (Hockey, 1995). The supervision relationship was found as often fraught and unsatisfactory as a result of neglect, abandonment, revenge of previous experience and disdain by the students (Johnson, Lee & Green, 2000). Due to this problem, Winfield report made some recommendations; training the supervisors should be introduced to the entire process of Ph.D., monitoring students' progress, showing more concern for the improvement and facilitating students' academic endeavours by providing more facilities and learning equipment.

On the other hand, Hockey (1995) ascribed the problem in supervision especially in social science to lack of training for supervisors and supervisees alike. According to him, training supervisors on how to supervise showed its efficiency and effectiveness in completion rates. Moreover, Chiang (2003) also emphasized on the role of training for both supervisors and research students. This training would give both supervisors and supervisees opportunities to interact successfully, which would enhance academic stimulation and promote collegiality, as well as reduce the segregation and, diminish level of loneliness. Supervisors need to undergo a series of training especially on how the students' psyche and emotion could be handled effectively. Moreover, Hockey (1995) emphasized that the content of training for supervisors should encompass both intellectual, organizational matters and pastoral skills. He further asserted that since experienced supervisors are rare in the universities today even in well-established high institutions, providing an effective training for the supervisors is a dire need. Swinnerton-Dyer (1982) demonstrated how science students were frequently associated with their supervisors, while the chance of often meeting with supervisors was not available for social science students. In addition to that, social sciences supervisors rarely attended any training courses to enhance their supervision ability but rather they merely relied on trail and error approach. Blume and Amsterdamka (1987) argued that the research process is totally different between natural sciences and human sciences in terms of academic culture between disciplines; when the former is laboratory-based study where things are more précised and accurately defined, the later is more complex, vague and dynamic. In addition to catering to intellectual development, both supervisors and research students should be familiar with different types of research methods not only the one that they use in their research (Chiang, 2003). It is an institutional responsibility to ensure that Ph.D. process is unique and reaches the expectation standard because the overwhelmingly majority of Ph.D. candidates return back to the institution as lecturers or academic surroundings (Chiang, 2003).

It is unjustifiable to attribute the supervisors' deficiencies as the sole factor for the students' academic deficiencies. However, supervision activities require the supervisors to develop a wide range of research relatedness and interpersonal skills. According to Wisker (2005) supervisors should be aligned with their practice and their learning behaviors with their supervisees. It is worth to note that the students are highly diverse in their academic ability, personality attributes, motivation and attitude (Zainal Abidin & West, 2007).

However, consideration of the supervision activities as the supervisors' private affair is one of the major obstacles that block supervision improvement. Supervisors are often reluctant to open their practice to criticism or benefit even from experienced supervisors. An open communication is crucial to identify the shortfall perceived by the students (Zainal Abidin & West, 2007). Factors mentioned include quality of supervision, part-time/full time status, financial situation of the candidate, and general environment. Generally, supervisors get assistance from the administration only on technical matters such as regulation and guidelines on supervision process without any concrete advice or training on how supervisors can enhance themselves especially on the body of knowledge and methodology in order to contribute significantly to the project and research (Pearson, 1996).

Supervisors should enhance themselves academically and professionally to be able to produce students with high standards and encourage them to be able to work independently. Although different supervisors have different style preference, they should put more efforts on appropriate ways to guide, direct, encourage, nurture the students' skills, foster of students' creativity, share ideas, and learn from

students' experiences and knowledge if necessary. According to Acker, Hills and Black (1994), "several students indicated that infrequent supervision had taught them how to be assertive, perhaps seeking help elsewhere, or to be better organized. Some thought that their supervisor had intended this outcome" (p.494). Another researcher (Holdaway, 1997) emphasized that some students were abandoned and left alone by their supervisors during their research activity period. He reported one Canadian research administrator said that "we must get humanities faculty members to feel that students are not a nuisance in their research" (p.67). Fostering supervisors' competence aims to integrate various needs and demands innate in the learning situation such as the supervisee's needs of professional enhancement, needs of educational development in all sense of the term, and personal needs of the supervisor for academic improvement. This personal needs of supervisor could be generated through reading, personalized reading (connecting it with own world), reflection, know-how, involving in research activities, and collegiality with their own supervisees (Linden, 1999; Elton & Pope, 1989; Burgess, 1997; Marsh, Rowe, Martin, 2003). Pearson (1996) emphasized that the quality of experience and the quality of outcome of a Ph.D. program depended largely on supervision and the individualistic nature of research.

In relation to supervision across gender, Booth and Satchell (1995) stated that women are more vulnerably to withdraw from a Ph.D. programme than men across all the subject areas. Men also completed faster in all subject areas. This finding indicated that even when supervision process accounts partially for burnout and late submission of Ph.D. project, male students are still benefiting compared to the female students. However, Wright and Cochrane (2000) differentiated between genders across their specializations. They found that females in science disciplines were slightly more successful to submit their dissertations as compared to humanities. However, in general the findings reinforced the previous studies that females performed less than men.

In Malaysia, especially the research universities have been trying to promote the postgraduate studies and enhance the quality of postgraduate research. The postgraduate students especially the Ph.D. students are considered as important resources for intensification of research and publication (Krauss & Ismail, 2010). Many workshops have been conducted for postgraduate students, irrespective of their specialization to train them for their future academic challenges. The universities are working hand-in-hand with various agencies to improve the postgraduate studies, and it is also providing necessary assistance and facilities for all postgraduate students based on the philosophy that research activities are very significant in the global knowledge economy and development. It is also firmly known that to achieve the universities objectives, the postgraduate students must involve in meaningful research to improve the effectiveness and efficiency of research supervision. Unfortunately, an inefficient Ph.D. supervision often leads to an increase in the time duration to obtain the Ph.D. degree; hence students must work far beyond their financially supported period. In other cases, supervision problems may force the candidate to leave his/her research career. Thus, supervision is a major issue when talking about PhD studies and needs to be considered seriously. Unfortunately, very few empirical studies on the quality of doctorial education are driven from students' perspective (Chiang, 2003).

Despite that many institutions have been criticized for malpractice of Ph.D. process of especially supervision process, the condition persist largely unchanged in many current circumstances. Thus, this study explored the Ph.D. students' views on the supervision practice at four public universities in Malaysia and their satisfaction level experiences across gender through different item functioning. In addition, the study also investigated the psychometric properties of supervision scale constructed by Van der Heide (1994), specifically the scale dimensionality; construct validity, endorsibility, and estimation of item and person score reliability of the scales.

2. Method

2.1. Participants

A sample of 153 Ph.D students from four selected universities, participated in this study. The researchers selected third year and graduated students for this study due to the fact that they were or had engaged in the supervisor-supervisee relationships with their respective supervisors during their writing processes. Thus, the students might objectively assess their supervisor performance, academic competency, characters and faculty supports.

2.2. Instrumentation

The instrument was adopted from previous study and used to assess PhD students' views on supervision process (Van der Heide, 1994). It was initially developed by Van der Heide, (1994) to measure the extent to which PhD and postgraduate research students have satisfactory experiences in relation to the quality of their research supervision. The instrument consisted of 49 items, with five distinctive factors (supervisors' academic competency, research

method competency, morals and characters towards supervisees, faculty supports, and supervisees' personal contribution to their research). The response categories were 7 for always true, 6 for almost always true, 5 for often true, 4 for sometimes true, 3 for seldom true, 2 for almost never true and 1 for never true. The internal consistency of the scale reported by Van der Heide, (1994) was ranged between .77 to .91. In the present study, the researchers tested internal consistencies of the adopted scale again employing the Cronbach's alpha and found it ranged between .96 to .98.

3. Results

Rasch model was then employed on collected data to psychometrically validate the used scale and its fulfillment of fundamental requirement to be considered sound scale to assess students' satisfaction towards supervision process. In addition to that, differential item functioning (DIF) was also used to identify items that function differently across gender.

Principal component analysis was used prior to the Rasch analysis calibration to identify the underlying factors (set of items loaded on a specific factor). The rationale behind using this technique was to fulfill the fundamental requirement of unidimensionality. After five meaningful and interpretable factors were extracted, Rasch calibrations were employed using the WINSTEPS program version 3.58.0 developed by Linacre (1991-2004). The Rasch analysis is based on assumption which derived from a basic probability that there is no relationship exists between a person's responses to different items after taking the ability into consideration (Pickard, Dalal & Bushnell, 2006). The correspondence between an individual's ability on a latent trait and the predicted response to an item is represented by an item characteristic curve which has Ogival (S-shaped) form. Item location along the continuum of the measure is expressed in log odds, or logits. Considering quality supervision, an expectation of the model for an item is that the probability of endorsing the item in the keyed direction increases as the amount of the quality supervision the individual holds increases. The Rasch analysis was selected for this study to identify the items that fit; that have equal items characteristics curve across gender. In order to prove reliability of the scale used for this analysis, the reliability of the overall and of each item were observed. Besides overall reliability of the model, WINSTEPS provides two mean square fit statistics; infit and outfit. The infit statistics "is an information-weight sum" and outfit "is based on conventional sum of squared standardized residuals" (Bond & Fox, 2001, p.176). Both infit and outfit are means square divided by their respective degrees of freedom, with an expected value of +1 and a range from 0 to positive infinity (Bond & Fox, 2001; Silver, Smith & Greene, 2001). The infit statistics are insensitive to unexpected responses to items far from a person's ability, while outfit is sensitive to unexpected ratings far from a person's ability. According to Silver et al. (2001), mean square statistics less than one (<1) suggested redundancy, dependency or constraint of data, while mean squares greater than one (>1) evidenced unexpected variability, inconsistency or extremism. Bond & Fox (2001) demonstrated that by saying infit e.g. 1.30 indicated 30% variation between the actual score and Rasch predicted score, while an outfit means square value of say 0.78 (1-0.22 = .78) showed 22% less variation in the observed score than modeled. Thus, the test of infit evaluates the consistency of item parameters across the person measured for each item. Data is combined across all items to provide an overall test of fit. On the other hand, the test of outfit shows the collective agreement for all items across persons. This is to support that item difficulties are consistent and stable (Waugh, 2001). Both item and person estimates allow researchers to determine how well an item measures a latent construct. It is worth noticing that the less variation between the actual score and the expected by the Rasch model is more desirable.

Furthermore, this study also assessed the different item functioning across gender. The Rasch model assumes that an additive structure underlies the observed data, the both participants, and items can be arrayed on a continuum, and that the items have equal discriminative power (Kan, Breteler, Van der Ven, & Zitma, 1998). Differential item functioning methods are widely used for detecting potentially comparison among quality supervision scale across gender. That means analyzing of individual differences in response tendencies as well as items discrimination (i.e. how well the item is able to discriminate between examinees holding different level of a latent construct). Thus, the Rasch model is capable to verify the fitness of the each item and person into the model spontaneously and provide the difficult level of endorsement (easy or difficult) of a person to each item. As obviously stated by Smith, ignoring or excluding misfit items from the analysis may not answer the complexity of the construct and may fail to provide an acceptable judgment. Thus, interpretation of extreme items may shed more light on how items were perceived and interpreted by an individual across the gender.

Although values range for both infit and outfit mean square fit statistics depends on testing situation and measurement purposes (Wright &Linacre, 1994), an acceptable range for this study is .60 to 1.40. The values within this range are considered relatively close enough to the perfect fit of the Rasch model. However,

the more the infit mean square and outfit mean square are further from an accepted range, the more other aspects are believed to play a role in determining the pattern of responses (Bond, 2001). Furthermore, in addition to infit and outfit means square, an index of reliability, and error estimation was also provided. The infit mean-squares are used to determine the fit of the item within the construct. Advancing average measures with each category and step calibrations ensure the rating scale measure is stable and accurate. Probability curves were used to visually inspect the rating scale category function.

3.1. Quality Supervision

Rasch model was applied to investigate Ph.D students' satisfaction of quality supervision, psychometrics properties of quality supervision scale and determine equivalence across gender. The satisfaction level of the students would be determine in term of how easy for the items involved can be endorsed. The easiness and hardness of an item would be expressed through the direction and magnitude of each item estimates, while negative sign of an item indicates easy to endorse (high satisfaction), the positive sign means hard to endorse (low satisfaction) (Bond & King, 2000). Precision of the estimates would be properly identified and accurately interpreted by adding and subtracting the measurement error from the magnitude of estimate.

Due to the preliminary nature of the study, a relatively broad fit criterion was used, and item with MSQ fit values greater than or equal to 1.40 were highlighted and explained. Referring to Table 1, the Calibration of the 49 quality supervision items yielded an acceptable model fit. Accordingly, in reference of Table 2, items separation reliability was .95, while the person separation reliability was .94, indicating high level of instrument consistency, items separation along the quality supervision scale continuum and that quality supervision estimates were well dispersed along the scale. The Standard Deviation (SD) of the item calibration was .39. The mean square infit ranged between .64 to 3.37, and the outfit mean square ranged between .60 to 4.54.

However, the rating scale (1-7) did not perform accordingly and the respondents haphazardly answered the questionnaires. It appears from the frequencies reported that respondents were not utilized the full range of the seven point scale, which was suspect in the overall analysis. Therefore, the researcher collapsed categories that did not act appropriately. Categories one and two were collapsed together (very strong disagree + strongly disagree) and categories six and seven (strongly agree + very strongly agree) also formed one scale. As a result, the item and person reliability slightly increased and the infit and outfit means square showed better fit.

Table 1: Summary of 49 measured items

	RAW		MODEL		INFIT		OUTFIT	
	SCORE	COUNT	MEASURE	ERROR	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	834.2	152.6	.00	.07	.99	-.5	1.07	-.1
S.D.	77.4	.6	.35	.00	.48	3.3	.71	3.5
MAX.	925.0	153.0	1.27	.08	2.86	9.9	4.00	9.9
MIN.	544.0	151.0	-.47	.06	.63	-3.9	.57	-3.7
REAL RMSE	.07	ADJ.SD	.34	SEPARATION	4.57	ITEM	RELIABILITY	.95
MODEL RMSE	.07	ADJ.SD	.34	SEPARATION	4.88	ITEM	RELIABILITY	.96

| S.E. OF ITEM MEAN = .05 |

Table2: Summary of 153 measured persons

	RAW		MODEL		INFIT		OUTFIT	
	SCORE	COUNT	MEASURE	ERROR	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	267.2	48.9	.71	.13	1.08	-.1	1.07	-.2
S.D.	46.4	.5	.68	.03	.45	2.5	.74	2.6
MAX.	335.0	49.0	2.12	.26	2.57	6.0	7.93	9.9
MIN.	113.0	46.0	-1.96	.11	.10	-9.2	.11	-9.0
REAL RMSE	.15	ADJ.SD	.66	SEPARATION	4.37	PERSON RELIABILITY		.95
MODEL RMSE	.13	ADJ.SD	.67	SEPARATION	4.99	PERSON RELIABILITY		.96

| S.E. OF PERSON MEAN = .06 |

Inspection of the infit and outfit mean square scores for 49 items revealed that six items were outside the set cutoff of 0.6 and 1.40. Both infit and outfit statistics of the six items were above the set cutoff, signifying high variability of response or misfit the model. Infit statistics of item 14 (*My supervisor has/had not taken over making decisions about my thesis)* was above the set cutoff (infit 1.62 and outfit 2.02) Item 15 (*My supervisor has/had not appreciated my work.)* (infit 2.84 and outfit 3.82), item 29 *(I have/had dealt with problems between my supervisor and me)* (infit 2.17 and outfit 2.73) item 45 *(*My *post graduate work (research) is like a hobby to me)* (infit 1.37 and outfit 2.01), item 46 (*Most of the time I have to force myself to do my post graduate study related work*) (infit 1.79 and outfit 2.04) *and item 47 (My post graduate study is pretty uninteresting)* (infit 2.58 and outfit 3.48). These infits and outfits suggested that the contents of the items were very hard for the respondents to endorse, in which indicating that they were less satisfied with them (Bond & King, 2000). It is obvious from the analysis that the contents of most of misfit items are negative in nature. With exception of these six items, the calibrations of the remained 43 items were within the set cutoff, indicating that the items were relatively closed enough to the perfect fit of the Rasch model.

The reliability of the items and persons was evaluated through using both the separation index (G) and reliability index (R). While the former is an estimation of how well people can be discriminated on the measured variable, the later which is conceptually similar to Cronbach's alpha is an indicator of the extent to which a different set of items measuring the same construct would reproduce the observed person scores (Bond & Fox, 2001; Hula, Doyle, McNeil, Mikolic,

2006). The examination of both separation index and reliability index yielded an acceptable level at 3.87 and R .96 for the persons while scored 4.55 and .96 for the items respectively.

Moreover, the low item estimates suggested easier endorsement or more satisfaction for the respondents. As it was shown in Table 3, the items estimation, more than half of the estimated values were negative while the positive estimate values for other items were very small in exception of departmental support items. This generally indicated that, Ph.D. students were satisfied with the supervision process's factors; supervisors' competency in both theory and methodology, attitudes towards supervisees, faculty support and practical outcomes of their research process. For example, estimate values table shows that respondents were more satisfied with the contents of item 19, "*My supervisor is/was supportive*", item 20 "*My supervisor is/was familiar with the field of research*", item 21, "*My supervisor respects/respected my ideas*", and item 37, "*My research have/had sharpened my analytical skills*" with estimate values *of* -.40, -.57, -.35, and -.41 with standard error ranged between .07 to .09 each respectively.

3.2. Different Items functioning across gender (DIF)

A differential item function within the Rasch analysis was employed to determine differences of respondents on quality supervision scale across gender. More precisely, to examine whether the scale is producing equivalent measures without discrimination in student satisfaction across gender that included 110 males and 43 females. The different item function occurs when different groups within the sample (e.g. males and females), despite equal levels of the underlying characteristic being measured, respond in a different manner to an individual item. Results in Table 4 revealed significant differential responses to only five items out of 49 items used in the study. The DIF contrast is the difference in difficulty of the item between the two groups, which is also significant for seven items at .62, p = .0155, -.55, p = .0007, .94, p =.0139, -.37, p = .0489, -.71, p = .0001, for item 7, 15, 20, 34, 47 respectively.

Table3: Quality Supervision of Ph.D. process input: 153 persons, 49 items measured

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
1	607	123	-.09	.10	.81	-1.3	.76	-1.6	.64	.58	50.4	44.7	1
2	641	126	-.29	.10	.60	-2.9	.65	-2.4	.68	.55	54.8	47.6	2
3	613	123	-.19	.10	.79	-1.4	.72	-1.9	.62	.57	50.4	44.9	3
4	633	132	.09	.09	.71	-2.3	.72	-2.1	.69	.60	43.2	42.1	4
5	665	133	-.15	.10	.77	-1.6	.74	-1.8	.68	.56	48.1	45.8	5
6	642	131	-.05	.09	.78	-1.6	.79	-1.5	.67	.58	45.0	43.2	6
7	662	132	-.16	.10	.65	-2.6	.65	-2.5	.68	.57	55.3	45.7	7
8	662	132	-.16	.10	.81	-1.3	.83	-1.1	.62	.57	52.3	46.1	8
9	645	131	-.04	.09	.67	-2.5	.70	-2.1	.70	.57	57.3	44.6	9
10	613	125	-.04	.10	.67	-2.4	.59	-3.0	.74	.57	51.2	43.8	10
11	676	135	-.14	.10	.85	-1.0	.77	-1.5	.58	.56	46.7	45.7	11

12	650	130	-.14	.10	.61	-2.9	.60	-2.9	.70	.56	53.1 45.5	12
13	618	123	-.20	.10	.96	-.2	.93	-.4	.58	.56	49.6 46.2	13
14	585	139	.70	.08	1.62	4.5	2.02	6.3	.42	.61	33.8 34.5	14
15	460	149	1.49	.07	2.84	9.9	3.82	9.9	.12	.65	11.5 31.0	15
16	604	122	-.08	.10	.77	-1.6	.88	-.7	.66	.58	45.9 44.8	16
17	592	117	-.25	.11	.65	-2.4	.68	-2.1	.67	.56	57.3 46.7	17
18	649	129	-.19	.10	.65	-2.6	.61	-2.8	.73	.56	52.7 45.8	18
19	625	121	-.39	.11	.72	-1.9	.72	-1.8	.67	.55	53.7 49.8	19
20	592	113	-.53	.12	.88	-.7	.85	-.8	.63	.53	61.1 50.3	20
21	647	125	-.38	.11	.84	-1.0	.79	-1.3	.65	.55	55.2 49.7	21
22	661	132	-.14	.10	.86	-.9	.78	-1.5	.65	.56	53.0 46.1	22
23	666	138	.12	.09	.95	-.3	1.06	.4	.66	.59	49.3 42.8	23
24	660	128	-.34	.11	.81	-1.2	.88	-.7	.62	.55	57.0 49.9	24
25	635	125	-.24	.10	.85	-.9	.90	-.5	.64	.55	52.0 47.8	25
26	685	134	-.29	.10	.68	-2.3	.65	-2.5	.67	.54	61.9 48.4	26
27	685	139	-.02	.09	.77	-1.7	.85	-1.0	.57	.57	48.2 44.4	27
28	661	128	-.36	.11	.80	-1.3	.98	-.1	.56	.55	51.6 50.2	28
29	578	136	.62	.08	2.17	7.4	2.73	9.2	.29	.63	33.8 35.3	29
30	636	125	-.24	.11	.85	-.9	.78	-1.4	.72	.57	56.0 48.4	30
31	684	132	-.38	.11	1.05	.4	1.03	.2	.55	.54	57.6 50.4	31
32	637	142	.44	.08	1.10	.8	1.08	.7	.60	.61	36.6 37.7	32
33	638	139	.31	.08	1.05	.4	1.00	.0	.60	.60	36.7 39.3	33
34	673	141	.14	.09	.89	-.8	.89	-.8	.63	.59	38.3 42.1	34
35	716	139	-.33	.10	.72	-2.0	.66	-2.4	.67	.55	61.2 49.4	35
36	663	128	-.39	.11	.89	-.6	.84	-1.0	.61	.55	61.7 50.0	36
37	642	123	-.48	.11	.69	-2.0	.74	-1.6	.62	.54	64.2 50.6	37
38	644	124	-.38	.11	.75	-1.6	.82	-1.1	.65	.55	62.9 50.7	38
39	667	131	-.24	.10	.83	-1.1	.83	-1.1	.62	.56	52.7 48.4	39

40	625	134	.25	.09	1.01	.1	.98	-.1	.63	.60	35.8 40.6	40
41	664	138	.11	.09	.95	-.3	1.08	.6	.57	.59	39.1 42.6	41
42	627	133	.19	.09	.94	-.4	.90	-.6	.64	.60	40.6 42.0	42
43	615	135	.33	.08	.87	-1.0	.83	-1.2	.66	.60	48.1 38.5	43
44	621	137	.36	.08	.88	-.9	.87	-.9	.66	.60	46.0 38.4	44
45	595	134	.41	.08	1.37	2.6	2.01	5.8	.42	.64	34.6 37.4	45
46	575	140	.71	.07	1.79	5.5	2.04	6.4	.36	.64	24.6 34.5	46
47	464	145	1.40	.07	2.58	9.9	3.48	9.9	.17	.63	16.8 29.8	47
48	687	139	-.03	.09	.85	-1.0	.92	-.5	.66	.59	50.7 44.8	48
49	664	129	-.32	.11	.80	-1.3	.74	-1.7	.70	.55	56.6 49.3	49
MEAN	633.7	131.4	.00	.10	.97	-.3	1.04	-.1			48.1 44.2	
S.D.	46.9	7.2	.42	.01	.46	2.9	.68	3.2			11.2 5.2	

Table 4: Supervisor and supervisee relationship

DIF class specification is: DIF=$S2W1

CLASS	PERSON MEASURE	S.E.	DIF CLASS	PERSON ITEM Number	DIF MEASURE	S.E. Name	DIF CONTRAST	S.E.	JOINT t	d.f.	Prob.
1	-.14	.12	2	1	.04	.18 1	-.18	.22	-.83	121	.4081
1	-.24	.12	2	2	-.44	.22 2	.19	.25	.79	124	.4325
1	-.19	.12	2	3	-.19	.19 3	.00	.23	-.01	121	.9952
1	.03	.11	2	4	.24	.16 4	-.21	.20	-1.10	130	.2746
1	-.10	.11	2	5	-.28	.20 5	.18	.23	.77	131	.4409
1	-.03	.11	2	6	-.09	.18 6	.07	.21	.31	129	.7597
1	-.01	.11	2	7	-.64	.23 7	.63	.26	2.45	130	.0155
1	-.05	.11	2	8	-.48	.22 8	.43	.24	1.76	130	.0810
1	.01	.11	2	9	-.22	.20 9	.23	.23	1.00	129	.3176

1	.03	.11	2	-.26	.21	.29	.24	1.20	123	.2311	
				10 10							
1	-.14	.11	2	-.13	.19	-.02	.22	-.07	133	.9455	
				11 11							
1	-.13	.11	2	-.15	.19	.02	.23	.08	128	.9368	
				12 12							
1	-.22	.12	2	-.15	.20	-.07	.24	-.32	121	.7529	
				13 13							
1	.71	.09	2	.67	.15	.04	.17	.24	137	.8102	
				14 14							
1	1.34	.08	2	1.89	.14	-.55	.16	-3.46	146	.0007	
				15 15							
1	.03	.11	2	-.40	.22	.43	.25	1.73	120	.0868	
				16 16							
1	-.26	.12	2	-.24	.21	-.02	.25	-.08	115	.9355	
				17 17							
1	-.13	.11	2	-.39	.21	.26	.24	1.08	127	.2833	
				18 18							
1	-.39	.13	2	-.35	.22	-.05	.26	-.18	119	.8588	
				19 19							
1	-.37	.13	2	-1.31	.35	.94	.38	2.50	111	.0139	
				20 20							
1	-.29	.12	2	-.70	.25	.41	.28	1.44	123	.1536	
				21 21							
1	-.11	.11	2	-.23	.20	.12	.23	.54	130	.5907	
				22 22							
1	.17	.10	2	-.03	.18	.20	.21	.94	136	.3466	
				23 23							
1	-.30	.12	2	-.48	.23	.18	.26	.71	126	.4819	
				24 24							
1	-.23	.12	2	-.27	.21	.03	.24	.14	123	.8891	
				25 25							
1	-.28	.12	2	-.28	.21	.00	.24	.00	132	.9978	
				26 26							
1	-.10	.11	2	.21	.17	-.31	.20	-1.52	137	.1299	
				27 27							
1	-.39	.13	2	-.27	.21	-.12	.25	-.50	126	.6210	
				28 28							
1	.52	.09	2	.84	.14	-.32	.17	-1.91	134	.0584	
				29 29							
1	-.26	.12	2	-.18	.20	-.08	.23	-.33	123	.7417	
				30 30							
1	-.34	.12	2	-.50	.24	.16	.26	.61	130	.5425	
				31 31							
1	.44	.09	2	.46	.15	-.03	.18	-.15	140	.8830	
				32 32							
1	.25	.10	2	.45	.15	-.20	.18	-1.13	137	.2616	
				33 33							
1	.03	.11	2	.40	.15	-.37	.18	-1.99	139	.0489	
				34 34							
1	-.39	.12	2	-.16	.19	-.23	.22	-1.05	137	.2962	
				35 35							
1	-.39	.13	2	-.37	.22	-.02	.25	-.07	126	.9430	
				36 36							
1	-.45	.13	2	-.58	.24	.13	.27	.48	121	.6316	
				37 37							

1	-.32	.13	2	-.53	.23	.20	.26	.78	122	.4397	
				38	38						
1	-.25	.12	2	-.19	.19	-.07	.23	-.29	129	.7721	
				39	39						
1	.34	.10	2	-.04	.18	.38	.21	1.84	132	.0682	
				40	40						
1	.17	.10	2	-.09	.19	.26	.21	1.23	136	.2191	
				41	41						
1	.20	.10	2	.19	.17	.01	.20	.05	131	.9587	
				42	42						
1	.43	.10	2	.06	.17	.37	.20	1.89	133	.0608	
				43	43						
1	.40	.09	2	.23	.17	.17	.19	.90	135	.3702	
				44	44						
1	.43	.10	2	.35	.16	.08	.19	.41	131	.6839	
				45	45						
1	.62	.09	2	.93	.14	-.31	.16	-1.90	136	.0593	
				46	46						
1	1.21	.08	2	1.92	.14	-.71	.16	-4.41	141	.0000	
				47	47						
1	-.02	.11	2	-.06	.19	.04	.22	.18	136	.8568	
				48	48						
1	-.31	.12	2	-.33	.22	.03	.25	.11	127	.9155	
				49	49						

The different item functioning occurred for item 7 "*My supervisor has/had given me constructive"* where it more easier for the male students to endorse (-.01) than their female counterparts (-.64), 15 "*My supervisor has/had not appreciated my work*" where it very hard for female to accept; low satisfaction (1.89) than their male counterparts (1.34). The different in response also found in items 20 "*My supervisor is/was familiar with the field of research*" where female respondents found it easier to accept (-1.31) than male respondents (-.37). Discrimination also occurred with item 34 "*I received appropriate assistance in locating a supervisor*", and item 47 "*My post graduate study is pretty uninteresting" in* where they were very difficult for females to endorse (low satisfaction) while were quite easy for male students to endorse (high satisfaction) .40, .03; 1.92, 1.21 females and males for both items respectively. This can be visually seen in the different item function plot displayed in Figure 1 as considerable differences across gender in the aforementioned items is obviously clear. (See figure 1)

4. Discussion

The Rasch model analysis was used to investigate students' level of satisfaction in the supervision process. The analysis was carried out firstly to validate the scale, secondly to test the satisfaction level of respondents and finally to assess the possibility of the items to function differently across gender. The results are largely in consistent with previous studies regarding students' satisfaction of supervision process (Hockey, 1991; 1995; Young, Fogarty, & McRae, 1987).

Interestingly, the present results generally support the construct and content validity of quality supervision scale (QSS) and provided evidence for the unidimensionality of the scale since the majority of the infit and outfit statistics of the items fall within an acceptable set cutoff of .60 – 1.40. It suggested for high level of scale consistency and item separation along the quality supervision scale continuum. However, the Rasch analysis revealed low level of satisfaction for faculty social and academic supports.

On the other hand, it was found that only 5 items of 49 were flagged for differential item functioning across gender. As also noted through the analysis and DIF plot, the Rasch-derived person scores are sufficiently precise to differentiate satisfaction level across gender. This finding should be treated with caution due to small sample size which can be interpreted as less representative of the population and also measure might also not mean the same thing across gender. Nevertheless, six items of QSS were fall out of set cutoff and they were highlighted and explained the possible reason for their misfitting.

As suggested by, Buttery, Rithcher and Filho (2005) the efficiency of Ph. D supervision arrangements depends on various stages of thesis lifecycle. Since different faculty has different supervisions styles, it is strongly recommend that future studies should also take to account quality supervision across different disciplines (Swinnerton-Dyer, 1982; Young, Fogarty, & McRae, 1987), as the researcher attempted to do initially, however, due to the lack of equal size across disciplines impede the objective. It is worth mentioning that was suggested that at least minimum of 30 respondents for each domain is required before different items function could be meaningfully employed (Linacre, 1994).

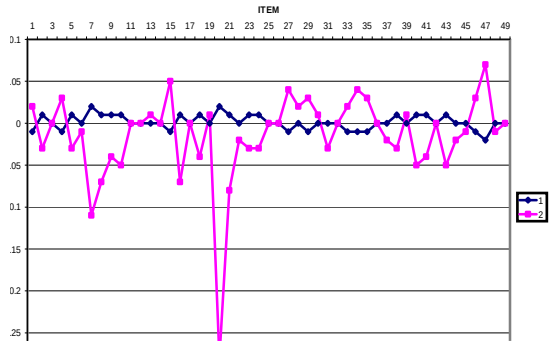

Figure 1: Person different item function plot

5. Conclusion

Generally, this study found that the Ph.D. candidates and graduates satisfied with the process of supervision especially with supervisors competent factors either in body of knowledge or in research methodology competency. They were also satisfied with their supervisors' moral interactions and their personal involvements and responsibilities towards their programs. However, although they study found general satisfaction of supervisees towards supervision process, the academic stage of the respondents is very significant element to direct their responses. More precisely, it was noted that respondents would naturally respond in positive tone and trivialize difficulty faced during their PhD process if they eventually and successfully completed their program irrespective of challenges. Thus, graduated study usually rated the supervision process positively even if otherwise happened.

Acknowledgement
We would to acknowledge the anonymous respondents from the selected public universities.

Corresponding Author:
Dr. Siti Aishah Hassan,
Department of Counselor Education & Counseling Psychology, Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
E-mail: siti_aishahh@putra.upm.edu.my, Siti8_hassan@yahoo.com

REFERENCES

1. Acker, S., Hills, T., & Black, E. (1994). Thesis Supervision in the Social Sciences: Managed or Negotiated? Higher Education, 28, 483-498.
2. Blume, S. & Amsterdamka, O.(1987). Post-Graduate Education in the 1980s. Paris, OECD.
3. Blume, S. (1995). Problem and Prospects of Research Training in the 1990s. In OECD (ed.), Research Training: Present and Future. Paris: Organization for Economic Co-operation and Development.
4. Bond, G & Fox, M. (2001). Applying the Rasch Model; Fundamental Measurement in Human Sciences. London: Lawrence Erlbaum Associates Publishers.
5. Bond, G. (2001). Client Satisfaction as an Educational Outcome: A Rasch Measurement Approach. Proceedings of the International Conference on Measurement and Evaluation in Education, Penang, Malaysia.
6. Bond, T. & King, J. (2001). Client Satisfaction as an Educational Outcome: a rasch measurement approach. Proceedings of the International Conference on Measurement And Evaluation in Education, Penang, Malaysia
7. Booth, A.L., Satchell, S.E. (1995). The Hazards of Doing a PhD: an Analysis of Completion and Withdrawal Rate of British PhD students in the 1980s. Journal of Royal Statistics Society Series A-Statistics in Society, 158, pp.145-159
8. Brew, A & Peseta, T. (2004). Changing Postgraduate Supervision Practice: a Programme to Encourage Learning Through Reflection and Feedback. Innovations in Education and Teaching International, 41(1), 5-22
9. Burgess, R. G. (1997). The Changing Context of Postgraduate Education in the United Kingdom. In Burgess, R.G. (ed.) Beyond the First Degree; Graduate Education, Lifelong Learning and Careers. Britain, Edmundsbury Press Ltd.
10. Burgess, R.G. (1994). Some Issues in Postgraduate Education and Training in the Social Sciences: an Introduction. In Burgess, R.G. (ed.) Postgraduate Education and Training in the Social Sciences: Processes and Products, London, Jessica Kindsley Publishers.
11. Buttery, E.A., Rithcher, E.M., & Filho, W.L. (2005). An Overview of the Elements that Influence

Efficiency in Postgraduate Superviosory Practice Arrangements. The International Journal Of Educational Management. 19, 1, p 7-26.
12. Chiang, K. (2003). Learning Experiences of Doctorial Students in UK Universities. The International Journal of Sociology and Social Policy, 23, 4-32
13. Earrleywine, M. (2006). Schizorypy, Marijuana, and Differential Item Functioning. Human Psychopharmacol, 21, 455-461
14. Elton, E. & Pope, E. & Pope, M. (1989). Research Supervision: the Value of Collegiality. Cambridge Journal of Education, 19(3), 269-275
15. Harris, M. (1996). Review of Postgraduate Education, Report. (Bristol, Higher Education Funding for England, Committee of Vice-Chancellors and principals and Standing Conference of Principals.
16. Hockey, J. (1991). The Social Science PhD: a Literature Review. Studies in Higher Education, 16(3), 319-332.
17. Hockey, J. (1995). Getting too Close: a Problem and Possible Solution in Social Science Ph.D. Supervision. British Journal of Guidance and Counseling, 23(2), 199-210.
18. Hula, W., Doyle, P. J., McNeil, M. R. Mikolic, J. M. (2006). Rasch Modeling of Revised Token Test Performance: Validity and Sensitivity of Change. Journal of Speech, Language, and Hearing Research, 49(1), 27-46
19. Johnson, L., Lee, A., & Green, B. (2000). The PhD and the Autonomous Self: Gender, Rationality and Postgraduate Pedagogy. Studies in Higher Education, 25(2), 135- 147.
20. Kan, C. C. Breteler, M. H., Van der Ven, A. H. & Zitma, F. G. (1998). An Evaluation of the DSM-III and ICD-10 Benzodiazepine Dependence Criteria using Rasch Modeling. Addiction, 93(3), 349-360
21. Krauss, S. E., & Ismail, I. A. (2010). PhD students' experiences of thesis supervision in Malaysia:Managing relationships in the midst of institutional change. *The Qualitative Report,15*(4), 802-822.
22. Linacre, J. M. (1994). Sample Size and Item Calibration Stability. Rasch measurement Transactions, 7, 328-333.
23. Linacre, J. M (2004). A User's Guide to Winsteps Ministep Rasch Model Computer Programs, Chicago.
24. Linden, L. (1999). The Contribution of Narrative to the Process of Supervising Ph.D. Students. Studies the Higher Education, 24(3), 351-369
25. Marsh, H.W., Rowe, K.J., Martin, A. (2002). PhD Students' Evaluations of Research Supervision. The Journal of Higher Education, 73(3),313-348.
26. Pickard, A. S., Dalal, M. R. & Bushnell, M. (2006). A Comparison of Depressive Symptoms in Stroke and Primary Care: Applying Rasch Model to Evaluate the Centre for Epidemiologic Studies. Depression Value in Health, 9(1), 1-7.
27. Seagram, B.C., Gould, J. & Pyke, S.W. (1998). An Investigation of Gender and Other Variables on Time to Completion of Doctoral Degrees, Research in Higher Education, 39(3), 319-335.
28. Sliver , B., Smith, V. & Greene, A. (2001). A Study Strategies Self-Efficacy Instrument for use with Community College Students. Educational and Psychological Measurement, 61(5), 849-865.
29. Van der Heide, G. (1994). Students and Supervision: the Review of Postgraduate Research Students in Education on Supervision. Paper presented at the Annual Conference of the Australian Association for Research in Education, Newcastle NSW, 28 November to 1 December.
30. Wisker, G. (2005). The Good Supervisor: Supervising Postgraduate and Undergraduate Research For Doctoral Theses and Dissertations. Macmillan, Palgrave.
31. Wright, T. & Cochrane, R. (2000). Factors Influencing Successful Submission of PhD Theses. Studies in Higher Education, 25(2), 181-195
32. Waugh, F. (2001). Measuring Ideal and Real Self-Concept on the Same Scale, Based on a Multifaceted, Hierarchical Model of Self-Concept. Educational and Psychological Measurement, 61(1), 85-101
33. Wright, B.D. & Linacre, J. M. (1994). Reasonable Mean-Square Fit Value. Rasch Measurement Transaction, 8(3), 370 -384
34. Young, K., Fogarty, M.P. & McRae (1987). The Management of Doctoral Studies in the Social Science. London, PSI.
35. Zainal Abidin, N. & West, M. (2007). Effective Meeting in Graduate Research Student Supervision. *Journal of Social Sciences* 3 (1), 27-35

2/1/2011