

Development of Allen and Meyer Commitment Scale Japanese Version: Based on Item Response Theory¹

Naotaka Watanabe and Koji Takahashi

Abstract

The objective of this study was to develop Japanese version of Allen and Meyer's three-dimension organizational commitment scale by item response theory (IRT). 18-items revised version was translated into Japanese and administered it to managers who worked for two electrical equipment manufacturing companies in Japan. Totally 849 responses were obtained and 844 data was proceeded to analysis after cleaning of the data. Through correlation analysis, it was found that there exists high correlation between affective commitment (AC) and normative commitment (NC) of which result is consistent with antecedent studies conducted in the U.S. The results of item parameters estimation followed by equating procedure using IRT revealed that there exist not a few items have aberrant parameter values. The result of differential item functioning (DIF) detection between two organizations showed that considerable number of items were suspected to function differently. Discussions were addressed about need for re-translation of the Japanese version and DIF detection between English and Japanese version in future studies.

Keyword

Organizational commitment, Three dimension model, Item response theory, Parameter estimation, Differential item functioning, Inter-language and intra-language DIF

Introduction

Organizational commitment² has gathered considerable attention both from researchers and practitioners. Organizational commitment research have dealt mainly with developing scales and using them since Kanter's (1968) convincing remarks on affective commitment, Becker's (1960) hypothetical argument on side-bet which sprouted continuous commitment later, and Etzioni's (1961) theoretical consideration on moral-based commitment which led to normative commitment. Because researchers and practitioners have regarded organizational commitment as an important variable in the organization settings, once valid and reliable scale was developed, they have concentrated on using it. A typical example is Porter et al.'s (1974) Organizational Commitment Questionnaire (OCQ) and its diffuse worldwide.

Clear was the tendency of concentrating on overwhelming scale instead of criticizing it and developing new ones. However, Allen and Meyer (1990) developed new scale on organizational commitment based on their three-dimension classification model, though OCQ was still popular device in commitment research. After it was published, it facilitated controversies on validity (including dimensionality) and reliability and it might take over the status once OCQ had. The need to translate Allen and Meyer's (1990) three-dimension commitment scale into another language is now apparent (cf. Ko, Price, & Mueller, 1997), but the translators and substantial users of translated version must solve some problems innate to translation.

The purpose of the present study was to develop the Japanese version of Allen and Meyer's (1990) three-dimension organizational commitment scale, to detect Differential Item Functioning (DIF) between similar but different samples from two Japanese companies, and to scrutinize the translated version by Item Response Theory (IRT) techniques

Three-dimension Model of Organizational Commitment

Comprehensiveness is the best description when we think of the definition of organizational commitment. Reflected on diversity of nature of organizational commitment, researchers employ various working definitions (Meyer & Allen, 1997; Takahashi, 1997a). Since Allen and Meyer (1990) and Meyer and Allen (1991) proposed three-dimesional model and scale of organizational commitment, researchers have tried to summarize past definitions consistent with three components (e.g., Dunham, Grube, & Casteneda, 1994; Racket, Bycio, & Hausdorf, 1994; Jaros, Jermier, Koehler, & Sincich, 1991; McGee & Ford, 1987; Meyer, Allen, & Gellatly, 1990; Meyer, Allen, & Smith, 1993). Three dimensions of organizational commitment hypothesized are described in Table 1.

Table 1 Three Dimensions of Organizational Commitment

Affective Commitment (AC)	Source: Attachment, Identification
<ol style="list-style-type: none"> 1. The relative strength of an individual's identification with and involvement in a particular organization (Steers, 1977, p. 46). 2. The attachment of an individual's fund of affectivity and emotion to the group (Kanter, 1968, p. 507). 3. An attitude or an orientation toward the organization which links or attaches the identity of the person to the organization (Sheldon, 1971, p. 143). 	

Continuance Commitment (CC) Source: Cost for Stay/Leave, Side-Bet

1. Commitment comes into being when a person, by making a side bet, links extraneous interests with a consistent line of activity (Becker, 1960, p. 32).
 2. Profit associated with continued participation and a "cost" associated with leaving (Kanter, 1968, p. 504)
 3. A structural phenomenon which occurs as a result of individual-organizational transactions and alterations in side bets or investments over time (Hrebiniak & Alutto, 1972, p. 556).
-

Normative Commitment (NC) Source: Obligation, Moral

1. The totality of internalized normative pressures to act in a way which meets organizational goals and interests (Wiener, 1982, p. 421).
 2. Socially accepted behaviors that exceed formal and/or normative expectations relevant to the object of commitment (Wiener & Gechman, 1977, p. 48).
 3. The committed employee considers it morally right to stay in the company, regardless of how much status enhancement or satisfaction the firm gives him or her over the years (Marsh & Mannari, 1977, p. 59).
-

Source: Meyer and Allen (1997), p. 12; Modification added.

Problems on Research Before Allen and Meyer (1990)

Organizational commitment has been a matter of researchers' and practitioners' interest for a couple of decades, which is expected to predict, concur, or result in job satisfaction, to increase productivity (e.g., Konovsky & Cropanzano, 1991; Nouri, 1994), to decrease turnover, tardiness, and absenteeism (e.g., Blau, 1986; Gellatly, 1995), and to indicate excellent culture of organizations (e.g., Guzzo & Noonan, 1994). Induced by active research situation and output, some researchers reviewed and summarized past research on commitment and tried to draw a reasonable conclusion on the relationships between commitment and other variables in 1980's (e.g., Mathieu & Zajac, 1985; Mowday, Porter, & Steers, 1982; Reichers, 1985).

As was revealed in Mathieu and Zajac's (1985) meta-analysis results, however, organizational commitment seems to correlate any possible indices which concern job-, work-, organization-, role-, demographic, and socio-cultural state/trait of employees. In a word, organizational commitment that is measured is a useful as antecedents, correlates, and consequences, but lacks decisiveness,

which totally conceal the essence of organizational commitment.

We must recall that organizational commitment had mostly been measured by Organizational Commitment Questionnaire (OCQ; Porter et al., 1974) until mid-1980's, thus the past research, summarized in review articles in those days, used OCQ as a device for measuring employees' commitment. If organizational commitment is criticized as "decisive," it is reasonable to attribute its origins whether to the commitment notion, per se, or to problems of OCQ's validity and/or reliability.

Presently, OCQ is criticized as less valid for four reasons (Takahashi, 1997a). First, OCQ's main span of measurement was only one dimension of organizational commitment (affective dimension), thus it could not cover and tap the breakdown or diversity of commitment notion into three sub-dimensions (e.g., Allen & Meyer, 1990; Meyer, Allen, & Gellatly, 1990; Meyer, Allen, & Smith, 1993; Meyer et al., 1989).

Second, because OCQ adopted a comprehensive definition of organizational commitment, it could not but include items to measure antecedents and consequences of commitment: that is, notional impurities. Porter et al. (1974) claimed that the nature of OCQ was characterized by at least three components: (1) a strong belief in and acceptance of the organization's goals and values; (2) a willingness to exert considerable effort on behalf of the organization; and (3) a strong desire to maintain membership in the organization. The sub-dimensions of OCQ may imply causal relationships between themselves. If we use OCQ for measurement, we cannot avoid the risk of measuring notional impurities out of organizational commitment itself.

Third, because OCQ prevailed and made the standard version of organizational commitment, many researchers employed it without considerable thought whether it was really valid for the measurement of commitment they intended to measure (Mathieu & Zajac, 1990; Reichers, 1985). This kind of misuse was apparent in operating scaling of sub-dimensions of OCQ. In some cases, all sub-scale scores were summed then treated unidimensional (as claimed in Sekimoto & Hanada, 1987), though the original study (Porter et al., 1974) did not recommend to do so.

Fourth and finally, cross-cultural research was not enough conducted to substantiate common ground of organizational commitment between English-speaking employees (mostly, employees in North America) and Japan using OCQ; that is, past research on OCQ implied that Japanese and North American employees committed to their organizations, based on the common ground.

It was not, however, tested or even well deliberated (Sekimoto & Hanada, 1985, 1986). One of few examples is Lincoln and Kalleberg's (1990) study. They conducted a cross-cultural research on organizational commitment of 4,567 workers of various companies in Indianapolis, U.S.A. and 3,735 counterparts in Atsugi, Japan, using OCQ with their own revision. They yielded that compared with U.S. respondents (Indianapolis'), Japanese respondents (Atsugi's) were less committed to the organization they join, because of (1) uniqueness of emergence of affection; (2) low rate of unemployment; (3) role and job ambiguity compared to the U.S. counterpart; and (4) low degree of family and life satisfaction (Lincoln & Kalleberg, 1990). Because the result did not support the broadly believed hypothesis, "Japanese employees are more committed to the organization than the U. S. counterparts," it became controversial and some different interpretations were presented. Currently, their interpretation of results is questioned in some studies (e.g., Fujimoto & Watanabe, 1998) from the viewpoint of psychometric procedure.

In summary, OCQ the most frequently used organizational commitment scale, is now questioned for its problems of notional impurities, content validity, limited range of commitment sub-scales measurement, and lack of universality of the model for cross-cultural comparison.

Detecting Differential Item Functioning (DIF) by Item Response Theory (IRT) ³

Until two decades ago, back-translation (BT) and pretest (Pre-T) using mass sample were regarded as a necessary and sufficient condition in translating scales to keep the translated version's validity and reliability as high as the original one has. We can propose that the shortcomings of BT and Pre-T is trade-off between literal and free translation (grammatical vs. content correspondence) and everlasting pursuit of non-biased samples. Because of the shortcomings, those techniques is now questioned for its usefulness in scale translation, influenced by diffusion of another powerful psychometric techniques, DIF detection using IRT, succeeded in conquering weak points of BT and Pre-T.

Back-translation is one of translation techniques that has been believed to be requisite and to qualitatively assure translation equivalence a priori between the original version and the translated one(s). However, when researchers execute BT and as the degree of difference between two languages compared with is greater such as English and Japanese, they hardly avoid the trade-off problems. When they employ literal translation style to keep equivalence in grammatical composure and form of the items between two languages, the degree of agreement with the items will be high, which is regarded as equivalent (success of BT); whereas if they adopt free translation style to maintain equivalence in the content of the items, the degree of agreement will be low, which is seen as less-equivalent (failure of BT).

Literal translation may make the expressions of translated items quite crude, artificial, and sometimes even strange. Higher agreement may secure the equivalence on expression (that is, similarity of expression) between the compared versions so that the researchers may not believe that the notion and content measured by the items is actually equivalent between the versions. Free translation may give translated items natural expression, which do not assure high agreement. In translating, this kind of trade-off is difficult to overcome. Besides, whichever style they take, researchers can agree with translation equivalence but cannot prove it, especially in the content domain (what is commonly measured between two versions of the same item), because BT is based on consent and agreement between translators and back-translators.

Pretest using mass sample cannot quantitatively assure the translation equivalence a posteriori because of possible fault of the assumptions on which it is based: Classical Test Theory (CTT). Of all assumptions that compose CTT, hypothetical population and non-biased sampling from the population is the most basic but controversial point in retaining translation equivalence. Because most of newly developed scales are still based only on CTT, scale developers must keep on unbiased sampling to gain translation equivalence between the two versions.

For example, when we translate a scale of depression developed in the U. S. in English language into Japanese and try to confirm translation equivalence between the two, we must conduct non-biased sampling from both hypothetical population, people in the U. S. and people in Japan. If the sampling was biased (for example, majority of the U. S. sample is "male, married, Protestant, lower-educated manufacturing workers in a rural area" and the Japanese counterpart is "female, non-married, Buddhist, highly-educated CPAs in an urban area"), can statistical parameters (e.g., Mean, SD) of both sub-population be comparable, or do we dare to conclude that translation equivalence is secured by comparing parameters? That is, pretest method based on CTT is sample-dependent in nature.

Therefore, we can scarcely conquer the problem of "non-biased sampling." A fundamental question, "Does the sub-population obtained from sampling truly represent population?", is left unsolved because it is difficult to prove. Statistical techniques to substantiate translation equivalence based on CTT require us to conduct sampling correctly and compare parameters measured by the original version and the translated one, but it is difficult. We must search for true representatives in two different settings; it is everlasting. As far as we cling to CTT, we will never be released from the dilemma of dual possible causal attribution: Is the statistical difference between means (or other possible parameters) of U. S. and Japan samples caused either by biased-sampling or by translation non-equivalence or by both? In other words, is the difference between samples caused because

they are really different, or is it because they are measured each by two similar but different instruments and statistically tested forcibly?

By using IRT techniques, we can gain "sample-free, sample independent" statistical parameter for item level. Hence, if an item of scale or test maintains the same meaning in hypothetical population, IRT statistics (item parameters: item discrimination, item difficulty, and pseudo-chance) calculated from any sub-population are identical. This nature of IRT connotes that identical item parameters are gained if the sample is believed to represent population and that the difference is attributed to inappropriate translation between the original version and the translated one (e.g., tapping different meaning between the same item of two versions). If there is difference between corresponding items of the original version and the translated one, we call it differential item functioning (DIF) between two corresponding items. The score of corresponding items with a clear DIF should not be compared, because there is a certain possibility of measuring and comparing the attitudes to different entities.

Inter-Language and Intra-Language Differential Item Functioning

Now that it is requisite to IRT-analyze pairwise items between the original version and the translated one in order to detect inter-language differential item functioning (IER-DIF): in this research, between English and Japanese version. By detecting IER-DIF before conducting cross-cultural research, we can avoid potential error of comparing scores of each item and scale (here, collected items) between original and translated versions which have DIF. It is useless to compare scores, if corresponding items and scales measure different entities.

Before trying to detect IER-DIF, however, we proclaim to confirm intra-language differential item functioning (IRA-DIF): DIF between different groups of participants in a single language. If IRA-DIF of item(s) is detected between two (or more) groups in a sole language, it cannot be denied that the IRA-DIF items measure different entities respectively, though they are completely identical in words expression. Besides, if IRA-DIF of item(s) is detected, users of translated version acknowledge the existence of IER-DIF between original and translated versions only after they conduct research and execute IRT analyses and they will not estimate the reason of IER-DIF occurrence. We, therefore, conduct IRT-analyses for detecting IRA-DIF, as a precedent step to conduct IER-DIF between English version and Japanese version.

Method

Targeted questionnaire and translation

In developing Japanese version, we decided to adopt revised version of three-dimension

commitment scale (cf. Meyer, Allen, & Smith, 1993) as a source because of three reasons. First, it gives respondents less burden in answering because the revised version composes 18 items, with 6 items for each dimensions of affective, continuous, and normative commitment. Second, it is actually "revised" thus reflects the notions of each component of commitment more accurately, based on empirical criticism furnished by research after the first publication of the original version. Finally, the revised version includes less "reversed" items. Reversed items in original version seem more straightforward and defiant (cf. Takahashi, 1994, 1997a), so we presumed that the revised one was more preferable for maintaining respondents' motivation to answer the questionnaire.

Translation was executed to fully reserve content correspondence by the authors, Watanabe and Takahashi, conferring with Takahashi's (1994, 1997a) attempt to translate the original version. We tried to relieve seemingly straightforward and defiant expressions of items whenever we thought they were there.

Sites and Participants

We conducted questionnaire survey for managers of two well-known electrical equipment manufacturing companies (hereafter, referred to as Company A and Company B, respectively). We obtained a total of 849 responses (Company A: n=432, Company B: n=412) and all was proceeded to further analyses after cleaning of the data. Questionnaire survey was executed from November till December, 1996.

All respondents except one participant were male (99.9%) and mostly married (96.0%),⁴ whose average age was 47.6 years old (SD=5.04) and average length of service (tenure) was 24.3 years (SD=6.68). Of all respondents, 73.8 percent earned BA/BS or higher degrees. Managerial rank depiction revealed that 21.1 percent were in "General Manager/Chief Director or Higher" (originally, Bucho Ijo) level; 75.2 percent in "Section Chief" (originally, Kacho) level; and 3.6 percent in "Lower than Section Chief" (originally, Shunin/Kakaricho). Comparing some demographic indices between Company A and B, Company B respondents' average age was significantly higher than Company A's (45.77 vs. 49.46, $t(842)=11.45$, $p<.01$; 22.75 vs. 25.87, $t(816.12)=6.96$, $p<.01$).

Measures

*Affective Commitment (6 items [including 3 reversed items]; $\alpha=0.69$).*⁵

Affective commitment (AC) means employees' affective (emotional) orientation toward the organization. The basis of AC may be senses of affection to and identification with the organization (Takahashi, 1997a).

Continuance Commitment (6 items; $\alpha=0.61$).

Continuance commitment (CC) indicates certain recognition of side-bet (unobtainable profit if withdrawing from the organization) and future potential unemployment possibility caused by the lack of low opportunity to gain new job elsewhere. The foundation of CC may be a rational calculation of these side-bet amount and possible loss in the future (Takahashi, 1997a).

Normative Commitment (6 items; $\alpha=0.77$).

Normative commitment (NC) represents moral and ethical recognition that it is their obligation to stay in the organization without quitting it. The base of NC may be the degree of employees to the organization (Takahashi, 1992), (socialization) pressures from organization (Takahashi, 1997a), and employees' personality (McKenna, 1994).

We adopted 7-point Likert scale for response scale of AC, CC, and NC and reversed the score of reversed items.

Overall Job Satisfaction (20 items; $\alpha=0.90$).

Overall job satisfaction (OJS) was measured by Short Version of Minnesota Satisfaction Questionnaire (MSQ; Weiss, Dawis, England, & Lofquist, 1967). MSQ Short Version was translated tentatively by the second author (cf. Takahashi, 1997b) and was adopted as a criterion of validity of AC, CC, and NC.

Analyses

First, we calculated basic statistics using SPSS Base for Windows statistical program (SPSS Inc., 1997). Then we executed t-test AC, CC, NC, and OJS scales between Company A and B. Finishing CTT-based statistical analysis, we IRT-analyze the data to detect IRA-DIF of items between Company A and B by BILOG3 (Mislevy & Bock, 1986). In conducting IRT analyses, response score for each item was dichotomized into 1 (positive; yes) or 0 (negative; no) with a cut-off point 4.01. Accordingly, central response ("Undecided" =4) was classified as negative response.

Item parameter estimation was executed using marginal likelihood method based on two-parameter logistic model for each item for Company A and B, respectively (cf. Baker, 1987, Takahashi, 1998); Watanabe, Noguchi, & Takahashi, 1994).⁶ After estimating each item parameters (a_i : item discrimination, b_i : item difficulty) for item sets of each company, we equated (or calibrated)

corresponding item sets - setting the identical origin of an item between companies - to enable DIF analysis afterwards.

Then we depicted item characteristic curve (ICC) of each item for each company⁷ and conducted DIF analysis by comparing a pair of ICC. In DIF detection, we employed "total area method" recommended by Lord (1980) and executed chi-square test for comparing areas between two ICCs of each item. If chi-square test for comparison between areas of two ICCs is statistically significant, it is implied that the item was biased thus DIF occurred.

Results

Descriptive Statistics and Correlation Coefficients

Table 2 describes descriptive statistics and Pearson's zero-order correlation coefficients among scales of overall, merged data. It is apparent that (1) the correlation coefficient between AC and NC is comparatively high ($r=.63, p<.01$); (2) the correlation coefficients between CC and AC and NC indicate similar values (.27 and .29, $p<.01$, respectively); (3) the correlation pattern of AC and NC with OJS is alike (.32 and .29, $p<.01$, respectively) whereas that of OJS and CC is not significant; and the correlation coefficients between length of service (tenure) and AC, CC, and NC resemble (.19, .13, and .20, $p<.01$, respectively).

Table 3 also depicts descriptive statistics and Person's zero-order correlation coefficients among scales of Company A data. Evidently we can read that (1) the correlation coefficient between AC and NC is comparatively high ($r=.63, p<.01$); (2) the correlation coefficients between CC and AC and NC indicate similar values (.22 and .29, $p<.01$, respectively); (3) the correlation pattern of AC and NC with OJS is alike (.34 and .31, $p<.01$, respectively) whereas that of OJS and CC is not significant; and the correlation coefficients between tenure and AC and NC resemble (.22 and .21, $p<.01$, respectively) while that between tenure and CC is not significant.

Table 4 also represents descriptive statistics and Pearson's zero-order correlation coefficients among scales of Company B data. Results acknowledge that (1) the correlation coefficient between AC and NC is comparatively high ($r=.62, p<.01$); (2) the correlation coefficients between CC and AC and NC indicate similar values (.32 and .29, $p<.01$, respectively); (3) the correlation pattern of AC and NC with OJS is alike (.30 and .28, $p<.01$, respectively) whereas that of OJS and CC is not significant; and the correlation coefficients between length of service (tenure) and AC, CC, and NC are alike (.15, .17, and .16, $p<.01$, respectively).

Table 2 Descriptive Statistics and Correlation of Scales (Overall)

	Mean	SD	1	2	3	4
1.AC	30.19	5.45				
2.CC	29.71	5.49	.27**			
3.NC	24.86	6.57	.63**	.29**		
4.OJS	58.35	8.81	.32**	-.05	.29**	
5.Tenure	24.29	6.68	.19**	.13**	.20**	.01

N=780; **p<.01

Table 3 Descriptive Statistics and Correlation of Scales (Company A)

	Mean	SD	1	2	3	4
1.AC	29.93	5.42				
2.CC	29.68	5.78	.22**			
3.NC	24.39	6.78	.63**	.29**		
4.OJS	58.12	8.54	.34**	-.05	.31**	
5.Tenure	22.75	6.02	.22**	.10	.21**	.06

N=398; **p<.01

Table 4 Descriptive Statistics and Correlation of Scales (Company B)

	Mean	SD	1	2	3	4
1.AC	30.47	5.47				
2.CC	29.74	5.18	.32**			
3.NC	25.35	6.31	.62**	.29**		
4.OJS	58.58	9.08	.30**	-.05	.28**	
5.Tenure	25.87	6.96	.15**	.17**	.16**	-.04

N=382; ** p<.01

Totally, analyses using aggregated, Company A, and Company B data produced similar results except the correlation pattern of OJS and CC which was not statistically significant. A remarkable point is high correlation between AC and NC and similar pattern of correlation was revealed between OJS and tenure and them; this is consistent with Hacket, Bycio, and Hausdorf's (1994) argument that AC and NC may influence each other.

t-Test of Scales

Results of t-test of score of each scale yielded that only NC is found statistically significant difference between Company A and B (24.39 vs. 25.35, $t(838)=-2.12$, $p<.05$), thus, NC of Company B was higher than that of Company A. No other significant difference was found of AC, CC, and OJS.

Estimating and Equating Item Parameters

Briefly, if there is considerable difference of value(s) of item discrimination and/or item difficulty for a single item between Company A and B, there may be IRA-DIF of the item between Company A and B. Table 5 to 7 explicate results of equated item parameter estimation of AC, CC, and NC, respectively, using Company A and B data. It is recognizable that AC and NC contains some items with "aberrant" value parameter(s): that is, $a_i < .50$ or $4.00 < |b_i|$ (see Roznowski, 1989, for detail). Whereas, CC composes less such items.

DIF Detection

We executed chi-square test in accordance with total area method, which Lord (1980) advocates, using estimated and equated parameters of each item between Company A and B. We depicted ICC of each item for Company A and B separately, then divided latent trait (θ) space (namely, horizontal axis of ICC) of each item into 41, and finally calculated the summation of value difference in each sub-division as a substitute of area difference. This summation denotes chi-square value of each comparison of ICCs and Table 5 to 7 indicate them of each item of AC, CC, and NC, respectively. Results of chi-square test were shown and considerable items were alarmed to detect IRA-DIF between Company A and B ($\chi^2 = 54.57$, $df=40$). Thus, all items of AC, all but for fourth (CC4) item of CC, and all except third (NC3) and fourth (NC4) of NC suggest IRA-DIF between Company A and B. In sum, Japanese version holds considerable number of items which are suspected to gain IRA-DIF.

Table 5 Results of IRA-DIF Analysis between Company A and B (AC)

Item	Discrim. (a_i)	Discrim. (b_i)	Difficulty (a_i)	Difficulty (b_i)	χ^2
	Company A	CompanyB	Company A	CompanyB	
1.	.622	5.733	.274	9.432	801.71*
2.	.432	4.541	.123	10.264	1002.58*
3(R).	.324	-2.533	.680	-.287	874.25*
4(R).	.217	-2.777	.522	-.424	779.57*

5(R).	.168	.007	.432	.313	610.47*
6.	.468	5.537	.171	9.555	719.96*
Mean	.372	1.751	.367	4.809	

Note: $\chi^2=54.57, df=40$; * $p<.05$

R: Reverse item

Table 6 Results of IRA-DIF Analysis between Company A and B (CC)

Item	Discrim. (a_i)	Discrim. (b_i)	Difficulty (a_i)	Difficulty (b_i)	χ^2
	Company A	CompanyB	Company A	CompanyB	
1.	.567	3.268	.833	3.412	215.67*
2.	.541	2.637	.438	2.891	170.79*
3.	.761	1.215	1.024	1.263	99.48*
4.	1.333	1.151	1.551	1.181	38.55
5.	.404	1.676	.423	2.131	165.90*
6.	.805	1.771	.689	2.122	178.53*
Mean	.735	1.953	.826	2.167	

Note: $\chi^2=54.57, df=40$; * $p<.05$

Table 7 Results of IRA-DIF Analysis between Company A and B (NC)

Item	Discrim. (a_i)	Discrim. (b_i)	Difficulty (a_i)	Difficulty (b_i)	χ^2
	Company A	CompanyB	Company A	CompanyB	
1.	.065	-5.309	.077	-8.051	374.42*
2.	.806	.795	.649	1.243	222.75*
3.	.853	-.197	.985	.083	44.97
4.	.571	2.593	.581	2.695	10.37
5.	.777	.564	.875	1.006	149.29*
6.	.843	1.195	.619	1.637	240.63*
Mean	.653	-.458	.631	-.231	

Note: $\chi^2=54.57, df=40$; * $p<.05$

Discussions

Correlation Patterns among Scales

The correlation pattern among AC, CC, and NC using total sample was consonant with the results of Meyer, Allen, and Smith (1986), using corresponding scale (three-dimension commitment scale,

revised version). Relatively high correlation coefficient was found between AC and NC, as past empirical research revealed (e.g., Racket, Bycio, & Hausdorf, 1994; Meyer, Allen, & Smith, 1993). As Meyer and Allen (1997, p.122) notes, AC and NC may not be independent each other, but in addition to this interpretation derived from empirical research using North American sample, we recall that the respondents of this study are Japanese managers. Also as Meyer and Allen (1997, p.108) argues by citing Hofstede's (1980) remark, "[NC] .. might be a better predictor than affective commitment in collectivist cultures that emphasize strong social ties (and obligations) and in cultures characterized by uncertainty avoidance where loyalty is considered a virtue... Japanese culture is more collectivist and uncertainty avoidant than is U.S. culture."

If AC and NC of typical Japanese employees are less differentiated in nature because of the result of successful socialization to Japanese collective society and organization, we can surmise that it is an important socialization task for employees (since they were newcomers/novices) to raise AC itself as well as NC, and the organization gives them certain pressure to do so: that is, AC is an index of successful socialization in Japanese organizations (cf. Takahashi, 1994).⁸ Japanese employees have to elevate their AC, as well as NC, as an indicator for their degree of organizational socialization. Japanese managers, as more socialized employees than their subordinates, must raise, maintain, and demonstrate their high AC and NC, which works as a target of modeling. Consistent with well heard comments from employees of Japanese organizations ("We, as a member of this company, must like our company, whatever the reason is."), AC and NC may be less distinct and they may be summed up altogether and treated as a significant scale of organizational socialization. We must deliberate if this might be the case with North American employees in the future study.

Correlation patterns between OJS and AC, CC, and NC were almost alike among overall, Company A, and Company B sample. OJS and CC did not correlate significantly whereas OJS and AC and NC did weakly, consistent with the results of Irving, Coleman, and Cooper (1997). Compared to the meta-analytic results of Mathieu and Zajac (1990, p. 178) between OJS and AC ($r=.688$) and OJS and CC ($r=.230$), this results implied greater correlation.⁹ We need more speculation whether OSJ and AC and CC correlate more weakly than other commitment scales.

Item Parameter Values

As shown in Table 5 to 7, item parameters of AC, CC, and NC of Company A and B indicate not a few aberrant values, $a_i < .50$ or $4.00 < |b_i|$ (Roznowski, 1989), for example, AC5 of Company A and B, CC5 of Company A and B, and NC1 of Company A and B. Because a pair of items is justified that there is not DIF between items if item parameter values of the pair are similar thus chi-square

value of the area difference gets smaller, we must distinguish between value aberrance itself and DIF detected by greater area difference.

Aberrant values of item parameters estimated and equated are found more in AC (all items) than in CC and NC. Thus AC itself is totally aberrant scale (see means of each item parameter estimated by each company sample), although CC and NC are not, as a scale, aberrant. Simply, because of high emergence of aberrant values of item parameters estimated by whichever sample, we need to revise the translation of each item of Japanese version and/or that of original version in the future. We can see so visibly that depiction of TCC (test characteristic curve) and TIC (test information curve) of AC yields less distinction between low- and high- θ sub-groups and less information as a whole, respectively.

Compared to AC, CC and NC revealed their superiority as a scale from the viewpoint of IRT. CC contains "difficult" items, which average respondents may find it difficult to say yes to, and it may discriminate relatively high- θ respondents sensitively. NC constitutes "a bit easy" items, which average respondents may find it easy to say yes to, and it may discriminate relatively low- θ respondents sensitively. By revising items with aberrant item parameter(s) in the future study, we can improve scale to make it more sensitive.

Recall that reverse items of AC (AC3, AC4, and AC5) show aberrant values, consistently, and that Japanese version includes only three reverse items of these. Now we must consider the reason that their item parameters were estimated aberrant: that is, in these reverse items, it may not be the case that negative response to reversed scale is positive.

Conspicuous IRA-DIF and Its Etiology: Need for Revising Japanese Version?

As is considered above, among items, IRA-DIF was detected conspicuously. Its etiology is now unclear but we would like to present "semantic difference" hypothesis: As each society has its primary culture and semantic system therein, an organization as a micro-society has its culture as a micro-culture. The content of a certain item (planned to measure common attitude within culture thus among micro-cultures) may function differently and stimulate distinct response among micro-societies (=organizations). This is only hypothetical consideration, we need to approach from the viewpoints of anthropology (e.g., Van Maanen & Barley, 1985), psychodynamic theory and psychoanalysis (e.g., Takahashi, 1997c), and post-modern philosophy (Markus, 1996).

Do we need to revise Japanese version? To answer this, we must collect more samples from different organizations as micro-society. Also, etic-emic dilemma among micro-societies should be

taken into consideration before trying to revise Japanese version.

Need for IRA-DIF Detection of Original Version

Results of IRA-DIF detection using Japanese version and sample implied there may be IRA-DIF in the original version. Further assessment is required to substantiate IRA-DIF in the original one, using IRT DIF detection techniques.

Need for IER-DIF between Original and Japanese Version: Toward a Cross-Cultural Comparison

After detecting IRA-DIF in original and Japanese version of three-dimension organizational commitment scale respectively and it is revealed that they are not, we may go into further analysis to detect IER-DIF between two versions. It is apparent that when we conduct translation of original version into another language which has different semantic system within, we must take a series of careful step: cautious translation, IRT-based parameter estimation and equating, item analysis, IRT-based IRA-DIF detection for translated version, and IRT-based IER-DIF detection between translated and original ones. Obeying this series of steps secures scale translation, which will envision successful cross-cultural research on validity of three-dimension organizational commitment model that was not yet substantiated.

Notes

1 This article is dedicated to the memory of my wonderful colleague as well as Ph.D. candidate, Mr. Koji Takahashi, who suddenly passed away. The original manuscript of this article, which had been still incomplete, happened to be found in the first author's hard disc drive in organizing the data files. The first author took over his work and completed it through modifying, editing and elaborating the original manuscript.

Portions of this study were presented in two continuous presentations by Koji Takahashi and Naotaka Watanabe, respectively, in Japanese, at the 13th Annual Conference, the Japanese Association of Industrial and Organizational Psychology, Hiroshima, Japan.

I gratefully acknowledge Professor John P. Meyer (Department of Psychology, University of Western Ontario, Canada) for giving us the permission of translating scale and developing Japanese version.

2 In other words, organizational commitment is referred to as "employee commitment," "workplace commitment," or simply, "commitment."

3 For details of IRT, see Crocker and Algina (1986), Hambleton (1989), Hambleton and Swaminathan (1985), Hulin, Drasgow, and Parsons (1983), and Lord (1980).

4 Separated, divorced, and deceased marriage was classified into unmarried category.

5 The explanations of three dimensions of organizational commitment are not necessarily

alike, compared to Meyer and Allen (1997).

6 Rationale for adopting two-parameter logistic model for item parameter estimation is that the respondents may not misunderstand content and context of each item, thus there appears slight chance for "pseudo-chance (guessing)" response, the third item parameter included in three-parameter model, to occur (see Takahashi, 1997a, 1998; Watanabe, Noguchi, & Takahashi, 1994 for details).

7 We used Ando's (1994) IRTLine for depicting ICCs.

8 We conducted additional exploratory factor analysis to acknowledge discriminant validity of AC and NC based on principal axis factoring and varimax rotation using whole data. Results revealed that AC and NC converged in one factor, composed of all items except NC2, NC3, and NC5 (scree test; first factor eigenvalue =4.24 and percentage of accounted variance=35.31%; whereas second factor eigenvalue=1.21 and percentage of accounted variance=10.08).

9 Mathieu and Zajac (1990) calls AC "Attitudinal Commitment" and names CC "Calculative Commitment."

Acknowledgement

This work was supported partly by research fund of Tokutei-Kadai-Kenkyu (Specific Academic Study), from Aichi-Shukutoku University (Project Number 18TT29).

References

- Allen, N. J., & Meyer, J.P. (1990). The measurement and antecedents of affective, continuance, and normative commitment to the organization. *Journal of Occupational Psychology*, 63, 1-18.
- Ando, M. (1994). *IRTLine: Program for drawing item characteristic and information curve and detection of differential item functioning* [Unpublished Computer Software].
- Baker, F. B. (1987). Methodology review: Item parameter estimation under the one-, two-, and three-parameter logistic models. *Applied Psychological Measurement*, 11, 111-142.
- Becker, H. S. (1960). Notes on the concept of commitment. *American Journal of Sociology*, 66, 32-42.
- Blau, G. (1986). Job involvement and organizational commitment as interactive predictors of tardiness and absenteeism. *Journal of Management*, 12, 577-584.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt, Brace, Javanovich.
- Dunham, R. B., Grube, J. A., & Casteneda, M. B. (1994). Organizational commitment: The utility of integrative definition. *Journal of Applied Psychology*, 79, 370-380.
- Etzioni, A. (1961). *A comparative analysis of complex organizations*. New York: Free Press.

- Fujimoto, T., & Watanabe, N. (1998, Mar.). The measurement equivalence of organizational commitment between the United States and Japan: A re-analysis of Lincoln and Kalleberg's data. *Paper Presented at the 1st Workshop, Japanese Association of Administrative Science, Nagoya, Japan.*
- Gellatly, I. R. (1995). Individual and group determinants of employee absenteeism: Test of a causal model. *Journal of Organizational Behavior*, 16, 469-485.
- Guzzo, R. A., & Noonan, K. A. (1994). Human resource practices as communications and the psychological contract. *Human Resource Management*, 33, 447-462.
- Hackett, R. D., Bycio, P., & Hausdorf, P.A. (1994). Further assessment of Meyer and Allen's (1991) model of organizational commitment. *Journal of Applied Psychology*, 79, 15-23.
- Hambleton, R. K. (1989). Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational measurement (3rd ed.)* (chap. 4).
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer Academic Publications.
- Hofstede, G. (1980). *Culture's consequences: International differences in Work-related values*. Beverly Hills, CA: Sage.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones- Irwin.
- Hrebiniak, L. C., & Alutto, J. A. (1972). Personal and role-related factors in the development of organizational commitment. *Administrative Science Quarterly*, 18, 555-572.
- Irving, P. G., Coleman, D. F., & Cooper, C. L. (1997). Further assessments of a three-component model of occupational commitment: Generalizability and difference across occupations. *Journal of Applied Psychology*, 82, 444-452.
- Jaros, S. J., Jermier, J.M., Koehler, J. W., & Sincich, T. (1991). Effects of calculative, affective and moral commitment on the turnover process: Evaluation of three structural equation models. *Proceedings of the 41st Annual Meetings of Academy of Management*, 220-224.
- Kanter, R. (1968). A study of commitment mechanisms in utopian societies. *American Sociological Review*, 33, 499-517.
- Konovsky, M.A., & Cropanzano, R. (1991). Perceived fairness of employee drug testing as a predictor of employee attitudes and job performance. *Journal of Applied Psychology*, 76, 698-707.
- Ko, J.-W., Price, J., & Mueller, C. W. (1997). Assessment of Meyer and Allen's three-component model of organizational commitment in South Korea. *Journal of Applied Psychology*, 82, 961-973.
- Lincoln, J. R., & Kalleberg, A. L. (1990). *Culture, control, and commitment: A study of work organization and work attitudes in the United States and Japan*. New York: Cambridge

University Press.

- Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.
- Markus, K. (1996). *Difference within organizational culture*. Ph.D. Dissertation, City University of New York, New York, U.S.A. (UMI #9618083)
- Mathieu, J.E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Journal of Applied Psychology*, 108, 171-194.
- McGee, G. W., & Ford, R. C. (1987). Two (or more?) dimensions of organizational commitment: Reexamination of the affective and continuance commitment scales. *Journal of Applied Psychology*, 72, 638-642.
- McKenna, E. (1994). *Business psychology & organisational behaviour*. Hove, East Sussex, U.K.: Lawrence Erlbaum Associates.
- Meyer, J.P., & Allen, N. J. (1991). A three-component of conceptualization of organizational commitment. *Human Resource Management Review*, 1, 61-89.
- Meyer, J.P., & Allen, N. J. (1997). *Commitment in the workplace: Theory, research, and application*. Thousand Oaks, CA: Sage.
- Meyer, J.P., Allen, N. J., & Gellatly, I. R. (1990). Affective and continuance commitment to the organization: Evaluation of measures and analysis of concurrent and time-lagged relations. *Journal of Applied Psychology*, 75, 710-720.
- Meyer, J.P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*, 78, 538-551.
- Meyer, J.P., Paunonen, S. V., Gellatly, I. R., Goffin, R. D., & Jackson, D. N. (1989). Organizational commitment and job performance: It's the nature of the commitments that counts. *Journal of Applied Psychology*, 74, 152-156.
- Mislevy, R. J., & Bock, R. D. (1986). *PC BILOG: Item analysis and test scoring with binary logistic models* (Computer Program). Mooresville, IN: Scientific Software.
- Mowday, R. T., Porter, L. M., & Steers, R. M. (1982). *Employee-organization linkages: The psychology of commitment, absenteeism, and turnover*. New York: Academic Press.
- Nouri, H. (1994). Using organizational commitment and job involvement to predict budgetary slack: A research note. *Accounting, Organization, and Society*, 19, 289-295.
- Porter, L. W., Steers, R. M., Mowday, R. T., & Boulian, P. V. (1974). Organizational commitment, job satisfaction, and turnover among psychiatric technicians. *Journal of Applied Psychology*, 59, 603-609.
- Reichers, A. E. (1985). A review and reconceptualization of organizational commitment.

- Academy of Management Review*, 10, 465-476.
- Roznowski, M. (1989). Examination of the measurement properties of the Job Descriptive Index with experimental items. *Journal of Applied Psychology*, 74, 805-814.
- Sekimoto, M., & Hanada, M. (1985). How to cope with the changing commitment of employees (I) : Research on commitment of 4,539 employees, 11 companies. [in Japanese] *Diamond Harvard Business*, November 1985, 84-96.
- Sekimoto, M., & Hanada, M. (1986). How to cope with the changing commitment of employees (II) : Research on commitment of 4,539 employees, 11 companies. [in Japanese] *Diamond Harvard Business*, January 1986, 53-62.
- Sekimoto, M., & Hanada, M. (1987). Structure of organizational commitment and its influences. [in Japanese] *Japanese Association of Industrial/Organizational Psychology Journal*, 1, 9-20.
- Sheldon, M. E. (1971). Investments and involvements as mechanisms producing commitment to the organization. *Administrative Science Quarterly*, 16, 143-150.
- SPSS Inc. (1997). *SPSS Base 7.5J for Windows* [Computer Program]. Tokyo: Author.
- Steers, R. M. (1977). Antecedents and outcomes of organizational commitment. *Administrative Science Quarterly*, 22, 46-56.
- Takahashi, K. (1994). A study on organizational socialization of newcomers: Examination of stage model of socialization. [in Japanese] *Unpublished Master's Thesis, Graduate School of Business Administration, Nanzan University, Nagoya, Japan*.
- Takahashi, K. (1997a). Item characteristics of organizational commitment scale and its application: Based on Allen and Meyer three- dimension commitment scale. [in Japanese] *Japanese Journal of Administrative Behavior*, 11, 123-136.
- Takahashi, K. (1997b). An analysis of Minnesota Satisfaction Questionnaire (MSQ) Short Version by Item Response Theory: Estimation of item parameters and its application to Human Resource Management. [in Japanese] *Journal of Law and Economics (Mie University)*, 15, 1-20.
- Takahashi, K. (1997c, May-June.). A tough day-one that lasts: Myths and facts of organizational entry in Japan. *Paper Presented at the 5th A. F. Jacobson Symposium in Communications (Symposium Title: "Workplace Socialization: Perspectives on Organizational Entry, Training, and Exit")*, Creighton University, Omaha, NE, U.S.A.
- Takahashi, K. (1998). Measurement of work attitudes (I): Job satisfaction. (in Japanese) In N.Watanabe & H. Noguchi (Eds.), *Item Response Theory: A New Approach toward Measurement of Organizational Psychology*. Tokyo: Hakuto Shobo.
- Van Maanen, J., & Barley, S. R. (1985). Cultural organization: Fragments of a theory. In P. J. Frost, L. F. Moore, M. R. Louis, C. C. Lundberg, & J. Martin (Eds.), *Organizational culture* (pp.

31-54.) Newbury Park, CA: Sage.

- Watanabe, N., Noguchi, H., & Takahashi, K. (1994). A development of Job Interest Index (JII) by utilizing Item Response Theory (IRT): Based on the "Company-Job" choice model in organizational entry. [in Japanese] *Japanese Association of Industrial/Organizational Psychology Journal*, 8, 13-27.
- Weiss, D. J., Dawis, R. V., England, G. W., & Lofquist, L. H. (1967). *Manual for the Minnesota Satisfaction Questionnaire (Minnesota Studies in Vocational Rehabilitation: Vol. XVI)*. Minneapolis, MN: University of Minnesota, Work Adjustment Project, Industrial Relations Center.
- Wiener, Y. (1982). Commitment in organizations: A normative view. *Academy of Management Review*, 7, 418-428.
- Wiener, Y., & Gechman, A. S. (1977). Commitment: A behavioral approach to job involvement. *Journal of Vocational Behavior*, 10, 47-52.