

The big five in the Japanese lexical approach

Yoshihiro MURAKAMI¹

In study 1, four psychology students collected personality trait words from a medium-sized dictionary under explicit rules, and investigated their appropriateness. As a result of three analyses, 539 nouns, 142 adjectives, 103 verbs, 37 adverbs, and 113 compound words were collected. The psychometric conditions of the big five were inferred from the old problem in semantic differential technique as follows: 1. The familiarity or frequency of the words must be high. 2. The variance of the rating must be large. In study 2, 370 university students rated themselves on 554 words. The 317 large variance words were analyzed, and the big five structure was obtained. The principal factor analysis and orthomax rotation were applied to the 100 representative words, and big five structure was re-confirmed. And the principal factor analyses and oblimin rotations were repeatedly applied to 20 words which belonged to the same factor. Three or two facets were obtained in each factor: Activity, seclusiveness and restraint facets in extroversion; envy, anger and selfishness facets in agreeableness; kindness, tenacity and orderliness facets in conscientiousness; energy and optimism facets in emotional stability; timidness, foolishness and weak-will facets in intelligence.

key words: big five, lexical study, variable selection procedure

The big five hypothesis is that the basic dimensions of personality consist of five orthogonal factors: extroversion (E), agreeableness(A), conscientiousness(C), emotional stability (N), and intelligence / openness for experience(O). This hypothesis goes back to the research by the lexical approach of Allport and Odbert(1936) who collected the 4504 terms designating personality traits from Webster's New International Dictionary.

First of all, Cattell(1943, 1945a, 1945b) examined the collected terms, and tried to reduce the size of the pool by grouping synonymous terms together. Thus, he reduced the pool to 35 bipolar personality clusters, and obtained 12 oblique factors as a result of personality rating. But Fiske(1949), Tupes and Christal(1961/1992), and Digman and Takemoto-Chock(1981) showed that Cattell's 35 trait terms could be described by the five orthogonal factors by reanalyzing his data and conducting the large-scale personality rating researches.

The big five hypothesis was also confirmed by Norman(1967) and Goldberg(1982, 1990, 1992, 1999). Norman(1967) collected 2800 personality trait terms from Webster's New International Dictionary (the third edition, 1961), omitted socially desirable or undesirable terms and vague terms which are inferior in descriptive power based on the the social desirability ratings and accuracy of self- and peer-ratings by the university students, and classified into 75 categories and 571 synonyms. Goldberg(1982, 1990) selected 1431 terms from Norman's 2800 synonyms by the accuracy rating with 8-point scale as a self-descriptor, analyzed 100 clusters derived from 339 terms, and obtained the big five structure, and Goldberg(1992, 1999) also repeatedly obtained the big five structure and developed big five markers.

The researches by the lexical approach were conducted predominantly in English but also in German, Dutch, Czech, Polish, Russian, Italian, Spanish, Hebrew, Hungarian, Turkish, Korean, and Tagalog (Philippines) language. It was to be noted that when terms that were

¹This article is a translated excerpt of Murakami(2002a, and 2002b). Correspondence should be addressed to Y. Murakami, Faculty of Education, Toyama University, Toyama, Japan. E-mail : murakami@edu.toyama-u.ac.jp

highly evaluative or describe physical features were included, usually structures with more than the five factors were obtained. The three factors — extroversion, agreeableness, and conscientiousness — were robust enough and replicated frequently, but other two factors emerged in unclear contents or in different styles. Presumably, these three factors were broader and were not easily influenced by the difference in the variable selection procedure than the other two factors. Thus, the big five hypothesis was only partially supported in these cross-cultural studies. But the replicability of the big five was affected by the variable selection procedure and by the kind of target (self, peer, semantic concept) that were rated (Saucier, Hampson and Goldberg, 2000).

In the lexical researches initiated by Allport et al. (1936), the neutral characteristic terms were generally collected; and terms which describe temporary mental states, feelings, social evaluations, physical features, and metaphors, etc. were excluded. This tradition has been succeeded by the research of Norman and Goldberg. Especially, it is common to exclude evaluative terms such as 'good—bad' in the methodology. A series of researches by Goldberg was based on the selected terms of Norman. When non-evaluative scales were used, the levels of self-peer agreement were as high as those for peer-peer agreement. Thus, non-evaluative scales offer superior generality between self- and peer- ratings (Saucier, Ostendorf and Peabody, 2001).

Saucier (1994) divided personality structure into descriptive and evaluative. He thought that the descriptive dimension will show a near-zero relation to desirability values of variables; an evaluative dimension will show a near-unity relation to desirability values of variables and near-zero relation to any descriptive dimensions. Saucier (1994) considered adjectives which had the mean value of 4.0–6.0 as neutral, based on the social desirability ratings of the terms by 100 university students. Beside this, there is a method of excluding the first factor which often reflects social desirable / undesirable terms.

This seems to be a reappearance of the old, unsettled problem of semantic differential (SD) technique in the field of the personality ratings as Saucier (1994) pointed out. For instance, Osgood, Suci and Tannenbaum (1957) proposed that the affective meaning system were composed of evaluation, potency, and activity factors. In summary, Osgood, May and Miron (1975) collected many data from all over the world and asserted the cross-cultural generality of this structure. However, these factors are not always obtained. They unite with other factors or divide into small pieces by the culture as well as the big five. Osgood, May and Miron (1975) made a list of adjectives (marker scales) which lead to the typical EPA structure, and when they are partialled out, peculiar factors to the culture were obtained.

This partialing-out technique was originally introduced by Kuusinen (1969). The reason for excluding marker scales was that the influence of evaluation, potency, and activity factor was too strong to obtain descriptive structure. However, he did not understand the psychometric meaning why the influence of marker scales was too strong. Nordenstreng (1969) discussed that the meaning measured by SD technique was divided into the hierarchy of the 'perceptual' and 'conceptual' concepts, although he did not propose a concrete technique to separate these concepts.

Murakami (1977a, b) thought that the influence of the partialing-out technique depended

on the scale variance, and acquired the data of Kuusinen(1969)², and obtained a mean value of marker scales and of other scales. In the former, 0.5686 and the latter was 0.8052 as expected. It seemed to be a significant difference, because the number of data was 25×100 . Thus, Murakami(1984) proposed the hierarchical factor structure in which the scales were divided into groups by the amount of information instead of scale variance. He also proposed a 3-level hierarchical structure: the basic structure of music was composed of warmth and intelligence factors; the affective structure was composed of mood, activity, familiarity and complexity factors, and no descriptive factors were obtained. However, no follower appeared because no computer program was readily available. Moreover, interest in a methodological problem of the SD technique was disappeared.

Let's return to the problem of social desirability concerning the big five structure. Generally, the evaluation-related terms have been excluded in the research by the lexical approach, and this dimension was most influential in the SD technique. The socially desirable terms are considered to be preferred by many people. Therefore, these terms must have small variance between individuals because the consensus of opinion on the terms will be high. Moreover, it is suggested that the mean values in the self- or peer- ratings by these terms will deviate from the center of scales.

The necessary conditions of introducing the big five are suggested as follows from the above-mentioned discussion.

1. The familiarity and frequency of the terms must be high. Because the validity and reliability of the self-rating and the peer-rating decrease when the vague or unfamiliar terms are used, it is natural to omit these terms.
2. The variances of the terms must be large, and the mean values of the rating must be located in the vicinity of the center of the scale.
3. The subjects must be composed of half men and half women. The men and the women are balanced in the majority of researches.

When attention is paid to such psychometric features, the terms can be selected only by the statistical procedure of the rating data. Thus the social desirability ratings such as Saucier(1994) are not necessary, and there is an advantage that the labor of the investigation can be saved.

LEXICAL RESEARCH IN JAPAN

Aoki (1971a) conducted the first lexical research in Japan. 6 investigators collected 3862 terms from the Shin-Meikai Japanese Language Dictionary(1988 edition) and grouped them into four categories; neutral words which showed personality trait or tendency; words related to evaluation; words which showed temporary feelings and body features; words which showed personality types. 455 words were selected from 517 words in the first category. The use frequency and social desirability of the words were examined, and they were classified into 10

²Kuusinen, J. Personal communication. 1975.

categories subjectively.

In Aoki (1971b), 80 university students and 80 workers rated 28 desirable items and 28 undesirable items from all categories of personality trait words, and the centroid factor analysis and varimax rotation were applied respectively. Three desirable factors from the former group, and three undesirable factors from the latter group were extracted. These factors might correspond to conscientiousness, agreeableness, and emotional stability in retrospect. The analysis of 64 items failed, because the memory capacity of the computer used at that time was very small.

Aoki (1972) added 125 ability words to 455 words, and tried an analysis similar to Cattell(1943) for 580 words. However, all words could not be analyzed at a time; so he divided subjectively into 11 groups, and analyzed the similarity ratings of 580 words by 109 male subjects, respectively. The results of factor analysis indicated 4 to 6 factors, and he selected 4 to 6 words in each groups. Thus, a 57 representative words list was obtained.

Aoki(1974) published "The dictionary of individuality expression" as a summary of his researches. The representative words were reviewed; the desirable words and the undesirable words were classified into another category; and finally he arranged the words to 14 categories.

However, after Aoki's researches, the lexical research was hardly conducted in Japan. It has been 30 years since Aoki's individuality expression words were obtained, and some words in the list are hardly used by the present university students to express personality traits. It is necessary to renew the lexical approach.

Recently, Tsuji(2001) collected 17158 personality trait words from "Koujien" (5th edition,1998) with the help of 3 university students, 3 postgraduates, and 17 psychology researchers. The 3 postgraduates and 15 psychology researchers rated 3-point scale of "Does a usual Japanese adult understand this word?" The mean value 1.5 or less was considered difficult, and 11145 words were chosen. In addition, the same 18 people rated on 3-point scale by "Whether the meaning was understood or not?" and "Whether they would use it as a personality trait word?" By the results of these three kinds of mean ratings, 3779 words were selected because their values exceed 2.5, and 400 basic words were selected because their value were all 3.0.

490 university students(54 males, 470 females) rated themselves on 5-point scale of these 400 words, and five factors rotated by equamax criterion. The first factor was extroversion(E), the second and the third were concerned with eeriness and aggressiveness, the forth was controlled intelligence(O), and the fifth was diligence and cooperation(uniting of C and A). Correspondence with the big five was not seen as a whole.

400 final words were chosen by the subjective judgments of the researchers. There was a possibility that the actual use frequency and familiarity of ordinary Japanese was different. Moreover, socially desirable words were not removed. Therefore, the range of standard deviation of the 5-point rating was widely distributed with 0.82-1.32, and the range of the mean values was also widely distributed with 1.77-4.06. Moreover, as most of the subjects were females, it is natural that the big five structure was not obtained.

PRELIMINARY ANALYSIS

The author started an independent lexical research almost at the simultaneous period as Tsuji(2001). The purpose was to collect basic personality trait words from 'Koujien'(CD-ROM, 5th edition,1998) as a preparation for the big five, and to investigate their appropriateness as personality trait words. 4 psychology students did the collection work based on the extraction rule. Provisional collection rules were drawn from Angleitner, Ostendorf and John(1990).

The final collection rules were as follows. Words shown by the bold-faced type were collected as personality trait words.

1. He is a character with the **noun**. ⇒ Example: He is a character with the **charm**.
2. He is a **adjective of noun origin** character. ⇒ Example: He is an **active** character.
3. He is a **adjective verb** character. ⇒ Example: He is an **oppressive** character.
4. He is a **adjective** character. ⇒ Example: He is a **gentle** character.
5. He is a character with **compound word**. ⇒ Example: He is a character with **small nature**.
6. He **verb** easily. ⇒ Example: He **angeres** easily.
7. He is a **conjugation of verb** character. ⇒ Example: He is a **distorted** character.
8. He is a **adverb** character. ⇒ Example: He is a **simple** character.

The following words were excluded: Words which showed temporary states and activities, words which showed social roles and interpersonal relationships, words which were difficult and seldom used, evaluation-related words, vague and metaphorical words, and geographical and occupational words. As a result, 950 words were collected.

The purpose of the second research was to investigate the understanding level of the words, and to make a more basic word list. 3 psychology students scrutinized the collected words and 14 improper words were deleted. About 300 words were allocated to a person, and 341 university students were instructed to exclude "Word whose meaning is not easily understood.", "Word hardly in use now", and "Word not related to the personality", to draw the line at the excluded words, and to blot them out.

The blotted-out percents of the subjects for each words were calculated. 587 words were below 10%. 165 words were of 10-20%. The words more than 30% decreased considerably. The lower bound of the blotted-out percent was assumed to be about 20% from these results, and 752 words were collected. However, the words such as "nastiness" and "nasty" which can be taken to have the same meaning were included in these 752 words.

After these analyses, 752 words were compared with Tsuji(2001)'s 400 basic words. Then, 226 words were common, and 174 words were different. Thus, 174 words were added, and 25 additional words were included, compared with Aoki(1971a)'s word list. In the third research, 125 university students blotted out Tsuji(2001)'s 174 words and Aoki(1971a)'s 25 words.

As a result of this analysis, 157 words from Tsuji(2001)'s list, 23 words from Aoki(1971a)'s list were added. Thus, 934 words in total were collected: 536 nouns(57.4%), 144 adjectives(15.4%), 110 verbs(11.8%), 37 adverbs(4.0%), 107 compound words(11.5%).

STUDY 1

As previously mentioned, Tsuji(2001) tried to obtain the big five structure in Japanese language, but no correspondent structure emerged. The reasons were as follows: The familiarity and frequency of the words were rated by the researchers, not by the students. The variances and mean values of the words were not controlled. (The socially desirable / undedireble words were not excluded.) And most of the subjects were females.

The purpose of this research was to extract the big five structure in Japanese based on the preliminary lexical analyses. A time restriction was placed on the rating work. About one hour was an upper limit during a lecture time at the university. And there are true-false, true-uncertain-false, 5-point, 7-point, and etc. scale forms. 7- or 5- point scale form is preferred to calculate the correlation coefficients because the values can be treated as continuous variables, but the working time increases. And the meaning of "definitely" in 7-point scale is different according to the individual. Thus the meaning of the answer is not clear as compared with true-false form(Kline, 1986). Then true-false form was selected because it required a shorter working time, and the words had to be reduced to about 500.

METHOD

Investigation Form

As a rule, words below the cross-out rate of 13% were collected from the basic word list. However, the words (for instance, carefree) which became an important key to the big five added more than the cross-out rate of 13%, and synonyms and antonyms were arranged. Thus 554 words were selected. However, because two words had been inadvertently included in the investigation form, 556 words were investigated. The rubies were shaken to all the words. The title of the form was "Personality Checklist".

Subjects

370 university students (150 male, average age 19.6 years old and $\sigma = 2.0$ and 220 female, average age 19.4 years old and $\sigma = 2.5$) answered validly to the mark card of MMPI-1, which was used as a substitution for the response sheet. 20 males and 25 females answered invalidly.

Method of Analysis

There is a relation $\sigma^2 = p_j(1 - p_j)$ between item variance σ^2 and approval rate p_j for true-false form(Shiba, 1972). When the approval rate is 0.5, the item variance is at its maximum value. Therefore, the words whose approval rate was in the neighbourhood of 0.5 in this analysis was retained for a factor analysis. The words that distorted the distribution were excluded from this work. It was expected that the big five structure would be obtained.

RESULTS

When each correlation of two words which were inadvertently included was examined, 0.781 and 0.786 were obtained. The numerical values corresponded to the test-retest reliability of this investigation. Considering the level of the present psychological testing, reliability was high enough though not as high as forecasted.

The approval rate of true or false was set to 0.650, and then the 237 words which exceeded this value were deleted, and the remaining 317 words were analyzed. When a false value was set for -1 and a true value for 1, the ranges of the average rating values of the words were distributed between -0.339 and 0.396, and the ranges of standard deviation were distributed between 0.915 and 1.00. The average rating values were in the vicinity of the center. Variances were considerably large because the maximal range of the score was 2.0.

SMC was placed in diagonal elements, and 30 factors were extracted. The eigenvalues turned out to be 32.29, 29.80, 16.03, 9.62, 6.51, 4.83, 3.94, 3.51, 3.30, 2.98, and 2.85.... The eigenvalues of the first five factors decreased rapidly, but those of the factors after sixth decreased slowly. Hence, it was judged that the factors latter sixth were an error factors.

The first factor was heavily loaded with "shy", "introvert", "quiet", "obedient", "cheerless", "quiet", "happy"(-), "uncommunicative", "keeping from", "withdrawn", "passive", "stand-offish", "active"(-), "vigor"(-), "no good at speaking" etc. This factor can be interpreted as extroversion(E) because it was related to introvert and extrovert contents.

The second factor was heavily loaded with "temperamental", "short-tempered", "to take a jaundice attitude", "to feel sick", "grudge", "to get mad", "to get angry", "selfish", "saucy", "persistent", "stroppy", "sourpuss", "boastful", "jealous" etc. This factor can be interpreted as agreeableness(A) because many of the words were related to sympathy or selfishness, and because the agreeableness(A-) dimension was related to cold, unkind, cruel, unsympathetic etc.(Murakami and Murakami, 2001).

The third factor was heavily loaded with "sincere", "conscientious", "faithful", "good intentions", "humanity", "kind", "earnest", "obedient", "gentle", "devoted", "warm", "cordial", "tenacious", "modest", "deliberate" etc. This factor can be interpreted as conscientious(C) because it was related to sincerity, conscience, cordiality.

The fourth factor was heavily loaded with "easygoing", "optimistic", "cool", "cheerful", "open", "happy", "frivolous", "active", "bold", "energetic", "suddern", "light", "positive", "cool-headed" etc. This factor can be interpreted as emotional stability(N) because it was related to easygoing contents.

The fifth factor was heavily loaded with "stupid", "half-finished", "discouraged", "sloppy", "rash", "scared", "overpowered", "strong-will", "thoughtless", "lazy person", "forgetful", "timid person", "bewildered", "person with no backbone", "childish" etc. This factor can be interpreted as openness for experience / intelligence(O) because it was related to non-intellectual contents, and scarce curiosity.

STUDY 2

The psychometric conditions of the big five were supported by the first analysis, but the consensus of opinions was not seen among researchers as for what facets exist and how to extract them. However, the exploration of facets is an important problem, because it influences construction methods of questionnaire.

To classify the 317 words in the first analysis according to the maximal factor loading into big five category, 45 words belonged to extroversion, 71 words to agreeableness, 67 words to conscienciousness, 66 words to emotional stability, and 68 words to intelligence. It took about 30 minutes to administer a 317 words check list in true-false form. Thus, the number of words were narrowed down in this analysis, and focused on the facets which relate closely to big five.

METHOD

Selection of words

20 words had been extracted from each factor on the factor loadings. However, for example, if the prefix was excluded, "jibunkatteña" and "kattena" became the same word (selfish), then "kattena" was deleted, since it had a smaller factor loading. And the words which did not satisfy simple structure — "omoshiroi (interesting)" and "maganuketa (stupid)" — were deleted, because the former was related to extroversion and emotional stability, and the latter to extroversion and intelligence.

Method of Analysis

The principal factor analysis and orthomax rotation (factor parsimony criterion) were applied to the correlation matrix of 100 words, and the big five structure was re-confirmed. And the principal factor analysis and oblimin rotation (biquartimin criterion) were applied to 20 words which belonged to the same factor.

RESULTS

Big five structure in 100 words

The result of factor analysis of 100 words was indicated in Table 1. It was almost the same factors as in the first analysis. No words were classified into a different category, however a little change emerged in factor loadings. That is, a more simple structure was obtained.

Table 1 The result of the orthomax rotation (factor parsimony criterion) of the 100 basic Japanese trait words

translated words	Japanese	E	A	C	N	O
shy	uchikina	0.665	0.034	-0.096	-0.206	-0.215
quiet	monoshizukana	0.625	-0.032	-0.177	0.000	0.024
lively	kappatsuna	-0.601	0.074	-0.206	0.413	0.153
happy	akarui	-0.597	-0.097	-0.255	0.318	0.072
obedient	otonashii	0.582	0.005	-0.194	0.001	-0.053
introvert	naikoutekina	0.580	0.100	-0.059	-0.241	-0.240
vigor	kakki ga aru	-0.578	0.072	-0.208	0.428	0.099
gay	nigiyakana	-0.578	0.152	-0.162	0.401	-0.069
keeping from	hikaemeno	0.572	-0.101	-0.274	-0.070	-0.068
outgoing	gaikoutekina	-0.551	0.008	-0.065	0.405	0.090
chatty	oshaberino	-0.549	0.133	-0.050	0.140	-0.161
talkative	hanashizukina	-0.546	0.103	-0.178	0.160	-0.029
cheerful	kaikatsuna	-0.540	0.071	-0.235	0.446	0.143
withdrawn	hikkomijianno	0.528	0.116	-0.077	-0.169	-0.282
active	katsudoutekina	-0.518	0.072	-0.200	0.391	0.132
cheerless	tsumaranai	0.498	0.193	0.046	-0.087	-0.307
passive	shoukyokutekina	0.493	0.120	-0.028	-0.298	-0.352
no good at speaking	kuchibetana	0.486	0.030	-0.070	-0.115	-0.199
uncommunicative	heisatekina	0.481	0.284	0.023	-0.149	-0.190
standoffish	yosoyososhii	0.456	0.207	-0.035	0.025	-0.233
temperamental	ikarippoi	-0.059	0.597	0.052	-0.111	-0.019
selfish	jibunkattena	0.045	0.551	0.246	0.194	-0.220
to get angry	haraga tatsu	-0.033	0.549	-0.023	-0.051	-0.034
grudge	netamu	0.017	0.539	-0.108	-0.132	-0.385
to feel sick	mukatsuku	0.104	0.539	-0.019	-0.009	-0.169
short-tempered	ki ga mijikai	-0.122	0.538	0.126	-0.116	-0.020
to get mad	atamani chi ga noboru	-0.075	0.535	-0.009	-0.019	0.015
stropy	hankoutekina	0.034	0.516	0.166	0.057	-0.017
saucy	namaikina	-0.020	0.515	0.072	0.224	-0.143
to take a jaundiced attitude	higamu	0.006	0.508	-0.159	-0.132	-0.359
persistent	shitsukoi	0.043	0.500	-0.069	0.071	-0.225
egocentric	jikochuushintekina	0.062	0.498	0.209	0.162	-0.176
sourpuss	hinekuremonono	0.165	0.492	0.048	-0.057	-0.349
vindictive	shuunenbukai	0.035	0.490	-0.189	0.047	-0.066
regretful	mirengamashii	0.077	0.477	-0.103	-0.082	-0.243

to have a sharp tongue	kuchiga warui	-0.104	0.475	0.106	0.033	-0.148
willfull	wagamamano	0.057	0.470	0.159	0.222	-0.172
jealous	shittobukai	-0.027	0.463	-0.125	-0.024	-0.180
to rival	hariau	-0.007	0.447	-0.127	0.240	0.150
boastful	tokuigena	-0.161	0.394	-0.152	0.218	0.010
sincere	seijitsuna	-0.042	-0.063	-0.678	0.189	0.085
faithful	chuujiitsuna	0.022	-0.042	-0.628	0.131	-0.031
conscientious	ryoushintekina	-0.167	-0.091	-0.625	0.182	0.014
kind	shinsetsuna	-0.156	-0.076	-0.616	0.195	0.116
good intentions	zenni ga aru	-0.148	-0.078	-0.609	0.163	-0.064
humanity	ninjou ga aru	-0.212	0.030	-0.590	0.100	-0.030
gentle	yasashii	-0.147	-0.151	-0.568	0.217	0.002
earnest	hitamukina	0.031	0.074	-0.565	0.179	0.173
obedient	juujunna	0.109	-0.032	-0.542	0.030	-0.163
devoted	kenshintekina	-0.028	-0.029	-0.513	0.145	-0.050
ardent	nesshinna	-0.131	0.150	-0.506	0.228	0.289
modest	kenkyona	0.277	-0.135	-0.505	0.032	-0.005
tenacious	nebarizuyoi	-0.006	0.112	-0.500	0.209	0.363
steady	kenjiitsuna	0.180	-0.007	-0.496	0.217	0.227
cordial	richigina	0.025	0.073	-0.493	0.113	0.099
merciful	nasakebukai	-0.045	0.136	-0.484	0.184	-0.008
brave	kenagena	-0.069	0.044	-0.476	0.252	-0.000
warm	ataatakai	-0.217	-0.137	-0.475	0.253	0.027
sense of responsibility	sekininkan ga aru	-0.163	0.065	-0.474	0.085	0.225
deliberate	nennirina	0.121	-0.028	-0.449	0.032	0.203
easygoing	kirakuna	0.039	-0.083	0.046	0.624	-0.182
optimistic	rakkantekina	-0.004	-0.036	0.027	0.615	-0.122
cool	heikina	-0.073	-0.122	-0.090	0.571	0.084
openhearted	kaihoutekina	-0.425	0.012	-0.044	0.529	0.057
cheerful	youkina	-0.437	-0.067	-0.233	0.511	-0.058
frivolous	noutenkina	0.012	0.041	0.039	0.504	-0.389
hedonism	kairakushugino	-0.046	0.052	0.165	0.480	-0.168
open	oopunna	-0.459	0.034	-0.045	0.478	0.033
suddernly	toppatsutekina	-0.010	0.278	-0.011	0.472	-0.029
light of heart	keikaina	-0.373	0.062	-0.160	0.471	0.128
calmly	heizentoshita	0.104	-0.015	-0.127	0.471	0.088
positive	maemukino	-0.241	-0.090	-0.189	0.467	0.267
bold	daitanna	-0.201	0.186	-0.089	0.465	0.136

The big five in the Japanese lexical approach

behavioral	koudoutekina	-0.406	0.079	-0.176	0.464	0.248
enjoyable	tanoshii	-0.375	-0.030	-0.270	0.462	0.060
unrestrained	kimamana	0.113	0.097	0.038	0.444	-0.118
energetic	enerugishuna	-0.424	0.120	-0.120	0.440	0.190
happy	shiawasena	-0.231	-0.050	-0.277	0.431	-0.021
pleasant	yukaina	-0.387	0.043	-0.285	0.428	-0.049
candid	kisakuna	-0.396	-0.012	-0.286	0.402	-0.066
discouraged	hekotareru	0.033	0.259	-0.044	-0.076	-0.585
half-finished	chuutohanpana	0.047	0.170	0.142	0.082	-0.576
to give up	akirameru	0.048	0.004	0.194	-0.086	-0.566
with no backbone	ikujinashino	0.198	0.142	-0.115	-0.189	-0.551
stupid	manukena	0.109	0.273	-0.144	0.152	-0.539
no reliance	tayorinai	0.299	0.101	0.064	-0.072	-0.533
timid person	shoushinmonono	0.208	0.099	-0.209	-0.196	-0.529
lazy person	namakemonono	0.149	0.250	0.142	0.050	-0.527
overpowered	megeru	0.077	0.376	-0.055	-0.184	-0.521
sloppy	darashinai	0.129	0.273	0.117	0.134	-0.516
scared	ojikezuku	0.113	0.217	-0.129	-0.100	-0.501
bewildered	urotaeru	0.127	0.213	-0.136	-0.153	-0.499
strong-will	ishi ga tsuyoi	-0.022	0.044	-0.228	0.181	0.476
rash	karuhazumina	0.028	0.361	0.088	0.243	-0.475
to throw out	nagedasu	0.111	0.230	0.160	0.035	-0.471
thoughtless	asahakana	0.037	0.241	0.056	0.075	-0.452
hasty	keisotsuna	0.024	0.267	0.157	0.262	-0.449
forgetful	wasureppoi	0.095	0.088	0.056	0.283	-0.434
childish	youchina	0.022	0.152	-0.031	0.201	-0.431
careless	fuchuuina	0.136	0.266	0.099	0.259	-0.427
contribution		7.610	7.443	7.745	6.772	8.629

contribution	7.610	7.443	7.745	6.772	8.629
--------------	-------	-------	-------	-------	-------

Facets in extroversion

The principal factor analysis was applied to 20 words in extroversion, the eigenvalues decreased to 7.456, 1.253, 0.665, 0.434, 0.232, 0.128 Because the accumulation of eigenvalue had become about 100% by the third factor, this seemed the upper bound of the number of factors. Then the 3 factors were rotated by oblimin method. The result was indicated in Table 2. The factors can be treated as an almost independent factor because the correlations among them were comparatively low.

Table 2 The extroversion items rotated by biquartimin criterion

translated words	Japanese	E_1	E_2	E_3
lively	kappatsuna	0.704	-0.132	-0.074
active	katsudoutekina	0.666	-0.072	-0.069
vigor	kakki ga aru	0.637	-0.049	0.052
cheerful	kaikatsuna	0.635	-0.086	0.002
gay	nigiyakana	0.485	0.103	0.335
happy	akarui	0.447	-0.154	0.115
outgoing	gaikoutekina	0.384	-0.188	0.138
uncommunicative	heisatekina	-0.099	0.499	0.049
withdrawn	hikkomizianno	-0.052	0.498	-0.100
introvert	naikoutekina	-0.191	0.489	-0.024
standoffish	yosoyososhii	0.029	0.483	-0.015
cheerless	tsumaranai	-0.057	0.477	-0.029
passive	shoukyokutekina	-0.159	0.456	-0.046
shy	uchikina	-0.142	0.382	-0.258
no good at speaking	kuchibetana	-0.048	0.319	-0.201
keeping from	hikaemeno	0.012	0.197	-0.420
obedient	otonashii	0.031	0.207	-0.419
chatty	oshaberino	0.185	0.029	0.417
quiet	monoshizukana	0.043	0.264	-0.377
talkative	hanashizukina	0.236	-0.006	0.363
contribution		2.523	1.906	1.080
		1.000		
corelation among factors		0.305	1.000	
		-0.298	0.316	1.000

The first factor was heavily loaded with “lively”, “active”, “vigor” etc. This factor was clearly related to activity, therefore it was named activity. This facet was frequently found, for example, as activity-adventurousness in Saucier, Hampson and Goldberg(2000), and activity facet in NEO-PI-R as a subcomponent of extroversion.

The second factor was heavily loaded with “uncommunicative”, “withdrawn”, “introvert” etc. This factor was named seclusiveness, because the word withdrawn was listed as a sociability subcomponent of extroversion, and these words were also classified as a parcel of extroversion in Saucier and Goldberg(2002).

The third factor was heavily loaded with “keeping from”, “obedient”, “chatty”, etc. This factor was named restraint, because the words such as talkative and reserved were included in a unrestrained subcomponent in Saucier et al.(2000). Restraint was enumerated as a different parcel to talkativeness in Saucier and Goldberg(2002). This factor can be interpreted as having fused two parcels.

Facets in agreeableness

The principal factor analysis was applied to 20 words in agreeableness, the eigenvalues decreased to 5.507, 0.972, 0.914, 0.498, 0.325, 0.227 Because the accumulation of eigenvalues had also become about 100% by the third factor, this seemed the upper bound of the number of factors. Then the 3 factors were rotated by oblimin method. The result was indicated in Table 3. The factors can be treated as an almost independent factor because the correlations among them were comparatively low.

Table 3 The agreeable items rotated by biquartimin criterion

translated words	Japanese	A ₁	A ₂	A ₃
grudge	netamu	-0.626	-0.017	-0.022
to take a jaundiced attitude	higamu	-0.575	-0.066	-0.051
regretful	mirengamashii	-0.542	0.027	0.015
sourpuss	hinekuremonono	-0.425	-0.033	0.151
jealous	shittobukai	-0.423	-0.042	0.091
persistent	shitsukoi	-0.394	-0.075	0.125
vindictive	shuunenbukai	-0.361	-0.079	0.072
to feel sick	mukatsuku	-0.286	-0.279	0.029
to rival	hariau	-0.180	-0.169	0.098
temperamental	ikarippoi	0.014	-0.633	0.039
to get mad	atamani chi ga noboru	-0.007	-0.592	-0.029
short-tempered	ki ga mijikai	-0.049	-0.496	0.029
to get angry	haraga tatsu	-0.164	-0.454	-0.024
to have a sharp tongue	kuchiga warui	-0.071	-0.302	0.207
stroppy	hankoutekina	-0.081	-0.284	0.214
selfish	jibunkattena	-0.025	-0.064	0.684
egocentric	jikochuushintekina	-0.063	0.037	0.664
willfull	wagamamano	-0.135	-0.061	0.421
saucy	namaikina	-0.166	-0.167	0.305
boastful	tokuigena	-0.170	-0.073	0.185
contribution		1.896	1.544	1.371
		1.000		
corelation among factors		-0.336	1.000	
		0.299	0.214	1.000

The first factor was heavily loaded with “grudge”, “to take a jaundiced attitude”, “regretful” etc. This factor was related to negative words concerning begrudging and envy, and seemed to be related to positive words such as obedience or generosity. But the latter words were deleted through variable selection procedure. This factor was named envy, because the word tolerant was listed in the agreeableness parcel in Saucier and Goldberg(2002). Moreover, Perugini & Di Blas(2002) listed revengeful as a agreeableness marker in the big five. On the

other hand, jealousy or envy was enumerated as emotional stability marker in Saucier and Goldberg(2002).

The second factor was heavily loaded with “temperamental”, “to get mad”, “short-tempered”, etc. This factor was named anger, because the word irritable was enumerated as a subcomponent of emotional stability in Saucier et al.(2000), and there was a facet of compliance concerning the control of aggressiveness in NEO-PI-R.

The third factor was heavily loaded with “selfish”, “egocentric”, “willfull” et al. This factor was named selfishness, because egoistic and egocentric were enumerated as a agreeableness marker in Perugini and Di Blas(2002) though no corresponding marker was found in Saucier et al.(2000) and Saucier and Goldberg(2002).

Facets in conscientiousness

The principal factor analysis was applied to 20 words in conscientiousness, the eigenvalues decreased to 6.520, 0.754, 0.623, 0.336, 0.279, 0.233 . . . Because the accumulation of eigenvalue had also become about 100% by the third factor, this seemed the upper bound of the number of factors. Then the 3 factors were rotated by oblimin method. The result was indicated in Table 4. The factors can be treated as an almost independent factor because the correlations among them were comparatively low.

Table 4 The conscientiousness items rotated by biquartimin criterion

translated words	Japanese	C_1	C_2	C_3
kind	shinsetsuna	-0.574	0.113	-0.059
gentle	yasashii	-0.534	-0.045	0.067
sincere	seijitsuna	-0.507	0.078	0.110
warm	ataakai	-0.505	-0.035	0.000
good intentions	zenni ga aru	-0.470	0.033	0.094
humanity	ninjou ga aru	-0.470	0.022	0.065
conscientious	ryoushintekina	-0.410	0.200	0.042
merciful	nasakebukai	-0.394	0.065	0.027
devoted	kenshintekina	-0.307	-0.009	0.222
sense of responsibility	sekininkan ga aru	-0.271	0.195	0.048
tenacious	nebarizuyoi	-0.108	0.565	-0.047
ardent	nesshinna	-0.141	0.528	-0.040
earnest	hitamukina	-0.175	0.409	0.055
deliberate	nennirina	0.071	0.319	0.262
obedient	juujunna	-0.101	-0.059	0.482
modest	kenkyona	-0.078	-0.026	0.429
faithful	chuuajitsuna	-0.202	0.070	0.385
steady	kenjitsuna	0.007	0.295	0.313

cordial	richigina	-0.125	0.125	0.280
brave	kenagena	-0.180	0.122	0.233
contribution		2.232	1.099	0.958
		1.000		
corelation among factors		0.344	1.000	
		0.368	-0.254	1.000

The first factor was heavily loaded with “kind”, “gentle”, “sincere” et al. This factor was named kindness, because kindness and agreeableness were classified into the same facet by this analysis, however the kindness parcel was separated from the agreeableness parcel in Saucier and Goldberg(2002).

The second factor was heavily loaded with “tenacious”, “ardent”, “earnest” etc. This factor was named tenacity, because it resembled perfectionism and meticulous parcels in Saucier and Goldberg(2002).

The third factor was heavily loaded with “obedient”, “modest”, “faithful” et al. This factor was named orderliness, because there was a marker of orderly in conscientiousness in Perugini and Di Blas(2002), however no correspondense parcel was in Saucier et al.

Facets in emotional stability

The principal factor analysis was applied to 20 words in emotional stability, the eigenvalues decreased to 5.963, 1.089, 0.497, 0.447, 0.349, 0.263 Because the accumulation of eigenvalue had also become about 100% by the third factor, this seemed the upper bounds of the number of factors. Then the 2 and 3 factors were rotated by oblimin method. The 3 factors solution resulted in high correlations among factors, so 2 factors solution was adopted, and was indicated in Table 4. The correlation between factors was a little large with 0.405.

Table 5 The emotional stability(neuroticism) items rotated by biquartimin criterion

translated words	Japanese	N_1	N_2
behavioral	koudoutekina	0.615	0.036
openhearted	kaihoutekina	0.585	-0.080
energetic	enerugishuna	0.584	0.028
open	oopunna	0.579	-0.046
enjoyable	tanoshii	0.566	-0.080
light	keikaina	0.552	-0.051
pleasant	yukaina	0.545	-0.085
cheerful	youkina	0.544	-0.182
positive	maemukino	0.487	-0.056
candid	kisakuna	0.484	-0.095

happy	shiawasena	0.453	-0.103
bold	daitanna	0.442	-0.091
cool	heikina	0.314	-0.312
easygoing	kirakuna	0.072	-0.560
optimistic	rakkantekina	0.103	-0.554
frivolous	noutenkina	0.016	-0.522
hedonism	kairakushugino	0.064	-0.429
unrestrained	kimamana	0.011	-0.394
suddenly	toppatsutekina	0.147	-0.335
calmly	heizentoshita	0.153	-0.329
contribution		3.650	1.641
corelation among factors		1.000	
		0.405	1.000

The first factor was heavily loaded with “behavioral”, “openhearted”, “energetic” and “enjoyable”, “light”, “pleasant”, etc. This factor was named energy, although there was no correspondent marker in the big five. It corresponded to an anxiety facet of NEO-PI-R and Saucier and Goldberg(2002) when interpreted as an indicator of no anxiety.

The second factor was heavily loaded with “easygoing”, “optimistic”, “frivolous” etc. This factor was named optimism, although there was no correspondent marker in the big five. It can also be interpreted as an indicator of no anxiety, but it differed from energy and pleasantness, because neurotic words were deleted by the variable selection.

Facets in intelligence

The principal factor analysis was applied to 20 words in intelligence, the eigenvalues decreased to 6.072, 0.831, 0.772, 0.374, 0.303, 0.135 ... Because the accumulation of eigenvalue had also become about 100% by the third factor, this seemed upper the bounds of the number of factors. Then the 3 factors were rotated by oblimin method. The result was indicated in Table 6. The factors can be treated as an almost independent factor because correlations among them were comparatively low.

Table 6 The intelligence (openness for experience) items rotated by biquartimin criterion

translated words	Japanese	O_1	O_2	O_3
timid person	shoushinmonono	0.536	-0.051	0.016
scared	ojikezuku	0.493	-0.085	0.057
with no backbone	ikujinashino	0.469	-0.033	0.148
bewildered	urotaeru	0.454	-0.145	0.005
overpowered	megeru	0.381	-0.085	0.216
discouraged	hekotareru	0.352	-0.071	0.271

hasty	keisotsuna	-0.045	-0.501	0.099
careless	fuchuuina	-0.001	-0.500	0.060
stupid	manukena	0.194	-0.492	-0.053
rash	karuhazumina	-0.045	-0.444	0.207
thoughtless	asahakana	0.121	-0.422	0.007
childish	youchina	0.128	-0.417	-0.086
forgetful	wasureppoi	0.065	-0.377	0.036
to give up	akirameru	0.128	0.072	0.515
to throw out	nagedasu	-0.033	-0.132	0.474
half-finished	chuutohanpana	-0.015	-0.204	0.465
strong-will	ishi ga tsuyoi	-0.094	-0.045	-0.412
lazy person	namakemonono	0.112	-0.173	0.371
sloppy	darashinai	0.004	-0.309	0.336
no reliance	tayorinai	0.201	-0.141	0.295
contribution		1.382	1.689	1.426
		1.000		
corelation among factors		0.332	1.000	
		-0.242	0.328	1.000

The first factor was heavily loaded with “timid person”, “scared”, “with no backbone” etc. This factor was named timidness, although there was no correspondent marker in the big five. Its positive meaning seemed to be related to boldness and eagerness.

The second factor was heavily loaded with “hasty”, “careless”, “stupid” etc. This factor was named foolishness, because its positive meaning corresponded to the marker of reflectiveness or intellectuality in Saucier and Goldberg(2002).

The third factor was heavily loaded with “to give up”, “to throw out”, “half-finished” etc. This factor was named weak-will, although there was no correspondent marker in the big five.

DISCUSSION

It was the first time that the big five structure was obtained in the Japanese lexical approach. The semantic content of the big five in the Japanese shows a high level agreement with its counterparts in Europe and America, though their nuance might be somewhat different. The three psychometric conditions of this research were shown to be appropriate to some degree, because the big five structure was obtained as expected. However, more words must be added in the future, because only 554 words were investigated in this research.

The number of factors increases unlimitedly if the personality trait words are added unlimitedly because the factor analysis is merely a technique to classify variables into several groups. It is incontrovertible that many factors exist beyond the big five(Saucier and Goldberg, 1998; Paunonen and Jackson, 2000). But the big five factors are stable and reliable, and it is important to consider what psychometric restrictions exist in obtaining these factors.

The facets in extroversion were named activity, seclusiveness and restraint. It was concluded that the three factors were included in extroversion, because these corresponded well to the big five markers. Neither shyness nor assertiveness were extracted, and talkativeness were included in seclusiveness. It is likely to be separated if a little more words were included.

The facets in agreeableness were named envy, anger and selfishness, although there was no correspondent marker in agreeableness. Envy was classified as a parcel of emotional stability in Saucier and Goldberg(2002). Warmth and sympathy parcels were not obtained. Only negative words were remained in agreeableness.

The facets in conscientiousness were named kindness, tenacity and orderliness. Tenacity might be a Japanese feature. However, kindness belonged to agreeableness in the big five of Europe and America. Existence of other facts was not proven by this analysis, though Saucier and Goldberg(2002) maintained that there were caution, organization, ambition and decisiveness parcels in conscientiousness.

The facets in emotional stability were named energy and optimism. The negative and neurotic words were not included. Therefore, the parcels of emotional excitability and fretfulness, etc. were not found. If both facets are interpreted as opposite to uneasiness, consistency with the big five can be seen.

The facets in intelligence were named timidness, foolishness and weak-will, although there is no correspondent marker in the big five. These will be named boldness, introspection and volition, when interpreted positively.

The big five in Japanese was similar to the structure in English as a whole, but its details were somewhat different from those of Europe and America. It is uncertain whether it is an artifact of this analysis or a reflection of Japanese original meaning system.

REFERENCES

- Allport, G.W. and Odbert, H.S. 1936 Trait names: A psycho-lexical study. *Psychological Monographs*, 47, (1, Whole No.211).
- Angleitner, A., Ostendorf, F. and John, O.P. 1990 Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality*, 4, 89-118.
- Aoki, T. 1971a A psycho-lexical study of personality trait words — selection, classification and desirability ratings of 455 words. *Japanese Journal of Psychology*, 42, 1-13(in Japanese with English summary).
- Aoki, T. 1971b A factor analytical study on personal desirabilities of personality trait words. *Japanese Journal of Psychology*, 42, 87-91(in Japanese with English summary).
- Aoki, T. 1972 A personality aspects based on multiple factor analysis on meaning-similarities among 580 personality trait words, *Japanese Journal of Psychology*, 43, 125-136(in Japanese with English summary).

- Aoki, T. 1974 *The dictionary of individuality expression*. Tokyo: Diamondsha (in Japanese).
- Cattell, R.B. 1943 The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476-508.
- Cattell, R.B. 1945a The description of personality: Principles and findings in a factor analysis. *American Journal of Psychology*, 58, 69-90.
- Cattell, R.B. 1945b The principal trait clusters for describing personality. *Psychological Bulletin*, 42, 129-161.
- Digman, J.M. and Takemoto-Chock, N.K. 1981 Factors in the natural language of personality: Re-analysis, comparison and interpretation of six major studies. *Multivariate Behavioral Research*, 16, 149-170.
- Fiske, D.W. 1949 Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology*, 44, 329-344.
- Goldberg, L. 1982 From Ace to Zombie: Some explorations in the language of personality. In C.D. Spielberger and J.N. Butcher (Eds.), *Advances in personality assessment* (Vol.1, 203-234). Hillsdale, NJ: Erlbaum.
- Goldberg, L. 1990 An alternative "Description of Personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Goldberg, L. 1992 The development of markers for the big-five factor structure. *Psychological Assessment*, 4, 26-42.
- Goldberg, L. 1999 A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I.J. Deary, F. De Fruyt and F. Ostendorf (Eds.) *Personality psychology in Europe*, vol.7. The Netherlands: Tilburg University Press. Pp.7-28.
- Kline, P. 1986 *A handbook of test construction: Introduction to psychometric design*. New York: Methuen and Co.
- Kuusinen, J. 1969 Affective and denotative structure of personality ratings. *Journal of Personality and Social Psychology*, 12, 181-188.
- Murakami, Y. 1977a A proposal of "stratum factor structure" on information measures. *The Japanese Journal of Psychology*, 47, 308-315 (in Japanese with English summary).
- Murakami, Y. 1977b On stratum factor structure: Containing the critics and theorizing of semantic differential method. *Psychologia—An international Journal of Psychology in the Orient—*, 20, 98-106.
- Murakami, Y. 1984 The stratified semantic structure of music: A proposal of 3-levels hierarchical model in semantic differential technique. *Japanese Psychological Research*, 26, 57-67.

- Murakami, Y. 2002a A collection of basic personality trait words. *The Japanese Journal of Personality*, 11, 35-49 (in Japanese with English summary).
- Murakami, Y. 2002b The big five structure in Japanese language and its psychometric conditions. *The Japanese Journal of Personality*, 12, (in Japanese with English summary, in print).
- Murakami, Y. and Murakami, C. 2001 *The handbook of the big five personality inventory*. Tokyo: Gakugeitoshō.
- Nordenstreng, K. 1969 *Toward quantification of meaning*. Helsinki: Suomalainen Tiedekatemia.
- Norman, W.T. 1967 *2, 800 personality trait descriptors: Normative operating characteristics for a university population*. Unpublished manuscript, Department of Psychology, University of Michigan (cited from Briggs, S.R. 1992 Assessing the five-factor model of personality description. *Journal of Personality*, 60, 253-293).
- Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. 1957 *The measurement of meaning*. Urbana: University of Illinois Press.
- Osgood, C.E., May, W.H. and Miron, M.S. 1975 *Cross-cultural universals of affective meaning*. Urbana: University of Illinois Press.
- Paunonen, S.V. 2000 What is beyond the big five? Plenty! *Journal of Personality*, 68, 821-835.
- Pergini, M. and Di Blas, L. 2002 The big five marker scales (BFMS) and the Italian AB5C taxonomy: Analyses from an etic-emic perspective. In B. de Raad and M. Perugini (Eds.) *Big five assessment*. Seattle: Hogrefe and Huber Publishers. Pp.281-304.
- Saucier, G. and Goldberg, L.R. 1998 What is beyond the big five? *Journal of Personality*, 66, 495-524.
- Saucier, G. 1994 Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology*, 66, 141-154.
- Saucier, G., Hampson, S.E. and Goldberg, L.R. 2000 Cross-language studies of lexical personality factors. In S.E. Hampson (Ed.) *Advances in personality psychology*, volume 1, Pp.1-36. London: Routledge. Pp.1-36.
- Saucier, G., Ostendorf, F. and Peabody, D. 2001 The non-evaluative circumplex of personality adjectives. *Journal of Personality*, 69, 537-582.
- Shiba S. 1972 Item analysis. In Hidano (Ed.) *Test 1 — The research method in psychology*. Volume 7. Tokyo: Toudaishuppankai. Pp.53-91 (in Japanese).
- Tsuji H. (Ed.) 2001 *Analysis of personality trait dimension based on the lexical approach in Japanese*. The report of grant-in-aid for scientific research (C) No.10610151 (in Japanese).

Tupes, E.C. and Christal, R. E. 1961 *Recurent personality factors based on trait ratings* (USAF ASD Technical Report No.61-97). Lackland Air Force Base, TX: US Air Force.(reprinted in *Journal of Personality*, 1992, 60, 225-251.)